NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

A PARAMETRIC COST MODEL FOR ESTIMATING ACQUISITION COSTS OF CONVENTIONAL U.S. NAVY SURFACE SHIPS

by

Kirk J. Loftus

September 1999

Thesis Advisor:	Timothy P. Anderson
Second Reader:	James N. Eagle

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1999	Master's Thesis

4. TITLE AND SUBTITLE
A PARAMETRIC COST MODEL FOR ESTIMATING ACQUISITION COSTS OF CONVENTIONAL U.S. NAVY SURFACE SHIPS

5. FUNDING NUMBERS

6. AUTHOR(S)
Loftus, Kirk J.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
None

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
When attempting to predict the acquisition costs of U.S. Navy surface ships, current models cannot produce a repeatable answer when the details of the acquisition program are not well defined. This thesis formulates a parametric model that predicts the average procurement cost of a conventional U.S. Navy surface ship based upon known (or assumed) physical and performance characteristics. The source data for the cost model is obtained from U.S. Weapons Systems Costs, a tabulation of annual procurement costs for major system programs, published by Data Search Associates. Standard regression techniques return cost estimating relationships able to predict average procurement cost from ship light displacement, ship overall length, ship propulsion shaft horsepower or number of propulsion engines. The formulated parametric cost model is approximate and appropriate only for rough order of magnitude studies, but can be used by the DoD cost community to produce justifiable estimates when other models do not have sufficient information to generate an answer.

14. SUBJECT TERMS
Parametric, Cost model, Surface ship, Acquisition, Estimation

15. NUMBER OF PAGES
108

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239.18

Approved for public release; distribution is unlimited.

A PARAMETRIC COST MODEL FOR ESTIMATING ACQUISITION COSTS OF CONVENTIONAL U.S. NAVY SURFACE SHIPS

Kirk J. Loftus
Lieutenant, United States Navy
B.S., United States Naval Academy, 1993

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1999

Author: ______________________
Kirk J. Loftus

Approved by: ______________________
Timothy P. Anderson, Thesis Advisor

____
James N. Eagle, Second Reader

____
Richard E. Rosenthal, Chairman
Department of Operations Research

ABSTRACT

When attempting to predict the acquisition costs of U.S. Navy surface ships, current models cannot produce a repeatable answer when the details of the acquisition program are not well defined. This thesis formulates a parametric model that predicts the average procurement cost of a conventional U.S. Navy surface ship based upon known (or assumed) physical and performance characteristics. The source data for the cost model is obtained from U.S. Weapons Systems Costs, a tabulation of annual procurement costs for major system programs, published by Data Search Associates. Standard regression techniques return cost estimating relationships able to predict average procurement cost from ship light displacement, ship overall length, ship propulsion shaft horsepower or number of propulsion engines. The formulated parametric cost model is approximate and appropriate only for rough order of magnitude studies, but can be used by the DoD cost community to produce justifiable estimates when other models do not have sufficient information to generate an answer.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS, ACRONYMS AND/OR ABBREVIATIONS

ρ	Coefficient of Correlation
α	Level of Significance
a	Adjusted (subscript)
AD	Destroyer Tender
AGOR	Oceanographic Research Ship
AGS	Oceanographic Research Ship
AO	Fleet Oiler
AOA	Analysis of Alternatives
AOE	Fast Combat Support Ship
APN	Aviation Program Navy
ARS	Salvage Ship
BCE	Baseline Cost Estimate
BEAM	Beam
BRAC	Base Realignment and Closure
BY	Budget Year
CAIG	Cost Analysis Improvements Group
CCA	Component Cost Analysis
CER	Cost Estimating Relationship
CG	Guided Missile Cruiser
Cov	Covariance
CV	Coefficient of Variation
cv	Cross-validated (subscript)
CVN	Aircraft Carrier, Nuclear
CY	Constant Year
DDG	Guided Missile Destroyer
DISP	Light Displacement
EA	Economic Analysis
ENGNUM	Number of Engines
FFG	Guided Missile Frigate
FSCA	Force Structure Cost Analysis
FY	Fiscal Year
ICE	Independent Cost Estimate
IID	Independent and Identically Distributed
LCAC	Landing Craft, Air Cushion
LEN	Length
LENBEAM	Ratio of Length to Beam
LHD	Amphibious Assault Ship
lm	Linear Model
LPD	Amphibious Assault Ship
LSD	Landing Ship Dock
MAX	Maximum Speed

MCM	Mine Countermeasures Ship
MHC	Coastal Minehunter
MSH	Hunting Minesweeper
MV	Multivariate Model
NCCA	Naval Center for Cost Analysis
NUM	Number
NWDC	Navy Warfare Development Command
O&S	Operations and Support
OLS	Ordinary Least Squares
PHM	Patrol Hydrofoil, Missile
POE	Program Office Estimate
PROP	Propulsion Type
R^2	Coefficient of Determination
ROM	Rough Order of Magnitude
ROM	Rough Order of Magnitude
RSE	Residual Squared Error
RSS	Residual Sum of Squares
SCN	Ship Construction Navy
SHP	Shaft Horsepower
SSBN	Ballistic Nuclear Submarine
SSN	Nuclear Attack Submarine
T1	Theoretical First Unit Cost
TAE	Ammunition Ship
TAGOS	Ocean Surveillance Ship
TAFS	Combat Stores Ship
TARC	Cable Repair Ship
TATF	Fleet Ocean Tug
Var	Variance
WBS	Work Breakdown Structure
WPN	Weapons Program Navy

EXECUTIVE SUMMARY

When evaluating new systems and strategies for programs with incomplete or loosely defined details, military decision makers have few tools with which to evaluate the expected program acquisition costs. Current tools have difficulty overcoming such limited information to produce a cost estimate. Robust methods such as cost extrapolation from a similar historical system or consulting with an expert to ascertain an opinion about the expected cost are difficult to validate. In addition, by its very name, any estimate will certainly be in error, and it is important to be able to determine the magnitude of that error.

This study utilizes parametric cost analysis to mitigate these problems, employing standard regression techniques to generate a number of parametric cost estimation models capable of transforming scant physical parameter data into a prediction of the procurement cost of a naval ship, including the uncertainty associated with that estimate. The model is also simple and sufficiently documented and may be used without specialized instruction.

The cost estimate produced by any of these models is justifiable as it has been based upon historical cost data. Ship procurement cost data are obtained from U.S. Weapons Systems Costs, published by Data Search Associates. Twenty-three surface ships, including small combatants, hydrofoils, cruisers, amphibious assault ships, oilers, support ships and others are included. Seven classes were removed as unsuitable: two ship classes were canceled before production or involved only the modification of existing ships; five additional ship classes were nuclear combatants and demonstrated

distinct cost and performance characteristics that made them unsuitable for inclusion in the database.

These models predict the average procurement cost (in constant 1999 dollars) of a conventional U.S. Naval surface ship. Four ship characteristics may be used as inputs: the ship light displacement, the ship overall length, the ship propulsion shaft horsepower, or the number of propulsion engines.

The models demonstrate a coefficient of variation (CV) between 74% and 83%, depending on the input variables selected; therefore predictions may still be expected to overestimate or underestimate the actual cost by more than 75 percent. The significant uncertainty of the model limits its applications to planning or evaluative purposes where a rough order of magnitude answer will suffice.

The models are unsuitable for applications requiring a tight tolerance around estimates; analysts seeking such predictions must select other methods. However, the models provide answers when no other tools are available. The Naval Center for Cost Analysis (NCCA) often requires rough order of magnitude estimates for a future ship's procurement cost. Similarly, the Office of the Chief of Naval Operations, Assessment Division (N81) requires models capable of estimating the costs of future systems to weigh against the benefits associated with particular strategic proposals. The models from this study are intended to provide these services.

The parametric acquisition cost estimating models are able to produce verifiable and defendable estimates from loosely defined parameters when detailed models can not. However, these models demonstrate significant limitations and would benefit from

additional refinements. New physical and performance data addressing weapons and sensor capabilities may capture aspects of procurement cost not addressed by the parameters chosen in this study. Also, an expanded database would further refine cost estimating relationships. However, within the scope defined herein, the models provide tools able to answer difficult questions about ship acquisition costs in a repeatable, defendable and justifiable manner.

ACKNOWLEDGMENT

The author would like to thank LCDR Timothy Anderson for his patient and pertinent advice, technical direction and oversight. His assistance throughout the project was indispensable. The author would also like to thank Dr. James Eagle for his continued support through all phases of this study, as well as Ms. Lisa Pfeiffer for her assistance and information.

Most of all, the author would like to thank God, his wife and family for their unfailing support, counsel, guidance and strength. Without their assistance this project would not have been possible.

I. INTRODUCTION

The Navy of today owes its shape to the strategies prevalent during the Cold War. During the course of that conflict, the Navy grew towards 600 ships, refining a strategy of sea control that re-cast the aircraft carrier from a significant component of the fleet into the preeminent instrument of naval strategy. Admiral Bernard Smith, commander of the Navy Warfare Development Command (NWDC) concurs; "Our force today was certainly designed around the open ocean and warfare that went along with the Cold War." (Peters) The strategy grew into the carrier battle group, defining both military and political strength, acting both as a symbol and instrument of foreign policy power.

And then it was over. The Berlin Wall fell; the Soviet economy collapsed; the United States emerged as the sole superpower. Without a clear opponent, military downsizing reduced the size of the Navy to 323 ships and left an acquisition plan that forecasts smaller numbers in the near future. The changing times and lack of clear direction foster uncertainty about what procurement plans the Navy should follow.

This does not imply that the Navy does not have a plan for the future. 'Joint Vision 2010' and 'From the Sea' provide the guidance for operational concepts that direct the strategy for fighting the fleet of today and stress capabilities that will be invaluable when fighting the fleet of tomorrow.

A. PROBLEM DESCRIPTION

Unfortunately, predictions in the face of uncertainty will never fit perfectly with the Navy of tomorrow, when tomorrow becomes today. To answer what should be done,

doctrines and strategies are corrected and updated to fit the changing world. However, the problem is not how to fight the forces in use or in production today. It is not what should be purchased to harness new technologies and tactical opportunities; although, it would appear to be so. The real problem is how to find the best economical solution.

The questions of today have already been answered. The innovations of 'dominant maneuver' and 'precision engagement' integral to the current strategy of converting information superiority into massed effects are well defined. (DoD Joint Warfighting Science and Technology Plan, chapter II.) Contractors provide competing proposals to fulfill the needs these strategies require. In the long term, however, the answers are harder to find.

New strategies in warfare, especially material decisions, must answer the questions: what are the benefits, and what will they cost? The first question may be answered by weighing anticipated capabilities of alternative systems against the perceived threats and challenges. Increasingly, they may be tested in simulated combat after making rudimentary assumptions about system capabilities. The second question—what will they cost—has fewer methods available to provide similar answers. In the absence of specific information, an analyst may either extrapolate from a single system that appears similar or consult with an expert to ascertain an opinion about the expected cost. Neither method offers much insight into the validity of the answer. The estimate is certain to be off; but how far off is anyone's guess. Parametric analysis offers an answer.

1. Selecting Parametric Analysis

Cost estimation may be divided into five distinct techniques, each with its own advantages and disadvantages. The first, *engineering estimation*, involves detailing every required item and process, assigning dollar figures to each identified element. Unfortunately, the details must all be known in order to pursue this technique, which is seldom the case when analyzing future warfighting strategies.

Another technique, *analogy estimation*, involves taking the known costs from a comparable system and stretching or twisting them until they appear similar to the unknown system. Although useful when the systems are similar, analogies do not provide significant information about the uncertainty of the estimation, because they are based on a single data point.

A third technique, *expert opinion,* harnesses the significant power of the human imagination to turn experience into an estimate, and often works well in the face of uncertain information. Unfortunately, one expert may produce a different estimate than another, and, without any means of substantiating one over the other, subject the estimate to human biases. There are methods, such as the Delphi and Consensus techniques, which combine different expert opinions into a single determination. (OA4702, p. 12-7, 12-8)

A fourth alternative, *extrapolation,* requires a well defined system both in place and already producing the product in question; estimates are generated by observing the actual costs of the existing system in the past and inferring that future costs will behave accordingly. While the technique applies well to predicting the cost of producing a few

more articles from a production facility currently in business, attempting to coerce estimates for new products represents a misuse of the technique. (NAWC, p. 9)

Finally, *parametric analysis* offers particular advantages when answering a cost analysis question of this type. Estimates may be produced at low cost, using a database of similar programs. The techniques also quantify the uncertainties associated with the cost estimate. Although limited by the quantity and quality of the database, the information may be updated easily, enabling the estimate to be reformulated quickly after a database addition. Finally, the parametric analysis provides a simple mathematical relationship that enables the user to quickly convert a set of independent variables into a reproducible cost estimate.

B. THE PURPOSE

The purpose of this study is to generate a series of parametric cost estimation models capable of transforming scant physical parameter data into predictions, including the uncertainty associated with the prediction, for the procurement cost of a naval ship. Cost estimates will be based upon historical cost data. The model must be simple enough to use with little instruction. Finally, it must be sufficiently documented, in order to allow similar techniques to be employed on new databases or subsets of data without excessive difficulty or repetition.

The models are intended to be high-level estimating tools, able to roughly estimate costs rather than precisely identify them. They are not intended for use by program managers to estimate current program costs. Rather, the models are designed to be used for evaluative purposes. The Naval Center for Cost Analysis (NCCA) often

requires Rough Order of Magnitude estimates for a future ship's procurement cost. Similarly, the Office of the Chief of Naval Operations, Assessment Division (N81) requires models capable of estimating the costs of future systems when weighing the benefits and risks associated with particular strategic proposals in Force Structure Cost Analysis. The models from this study are intended to provide these services.

THIS PAGE INTENTIONALLY LEFT BLANK

II. PARAMETRIC COST ESTIMATION PROCESS

The methodology for performing parametric cost estimation is well defined. The separation of the tasks into particular stages differs with background sources and presentation format, but the underlying milestones and their order are reasonably consistent. Section II will provide general information about how to conduct cost estimates and will provide an overview concerning the particular techniques this study will use in later sections.

The parametric cost estimating process begins when an analyst poses a question about the costs of an unknown system and collects information in the pursuit of an answer. It continues through the postulation and verification of cost estimating relationships (CERs) that describe system costs as mathematical functions of physical parameters and characteristics and ends when a useful model has been developed to answer that question.

Each step in the process outlines particular tasks that should be completed before progressing to the next step. By following the steps in order, the necessary foundations of the cost analysis will be completed before elaborate and logically risky conclusions are drawn. This cost estimating process is illustrated in Figure 1.

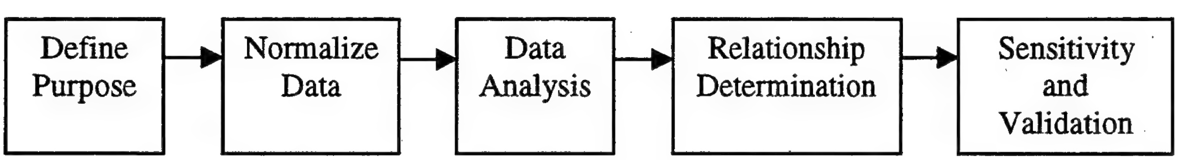


Figure 1: Parametric Cost Estimating Process (OA4702, p. 2-2).

A. DEFINING THE PURPOSE

As a first step, the analyst must determine the purpose of the cost estimate. This purpose will determine practically every aspect of the eventual analysis, including the time needed to complete it, the desired accuracy and precision of the study, appropriate analysis methods and the scope of the required data. (OA4702, p. 2-4, 2-5)

1. Cost Analysis Applications

Several specific types of cost estimates deserve additional attention, as they shall be specifically addressed in this study. The purpose of the cost estimate dictates the type of analysis performed.

a. Rough Order of Magnitude (ROM) Estimate

By their title, ROM estimates value a quick answer over a precise solution. Because an answer is available quickly, ROM estimates are able to approximate a funding requirement in advance of a detailed study, although the actual costs may be difficult to justify under scrutiny. A ROM estimate may be used in other applications, especially when comparing alternatives in the distant future. (OA4702, p. 2-8)

b. Feasibility Study

When a new concept begins evolving into a program, it invokes questions about whether the concept is attainable and practical. These questions may be addressed using a feasibility study to decide whether the new idea appears worth, in benefits and performance, the investment of time and money the program would require. Because the new program may not yet be well defined, feasibility studies may also include ROM estimates in their analysis. (OA4702, p. 2-8)

c. Economic Analyses (EA) and Analysis of Alternatives (AOA)

Economic Analyses compare two or more alternative investment decisions in terms of their costs and benefits. An Analysis of Alternatives is a specific form of EA used to compare alternative weapons systems in terms of their costs and effectiveness in meeting particular mission areas. (DoDINST 5000.2R, 2.4.1) Both are intended to aid decision makers in judging whether or not any of the proposed alternatives to an existing system or investment offer sufficient military and/or economic benefit to justify the cost. The analysis must be quantitative, specifying requirements, necessary performance criteria and particular means of evaluating the criteria. (DoDINST 5000.2R, 2.4)

d. Force Structure Cost Analysis (FSCA)

A force structure cost analysis addresses the cost of an entire concept or strategy. Instead of concentrating on a particular acquisition program, an FSCA evaluates the effects on cost of a change in the existing force structure. Examples include Base Realignment and Closure (BRAC) studies, downsizing the service strength of the Navy, embracing a new strategy of massed power projection; all raise questions about the cost of the changes. (OA4702, p. 2-17) An example FSCA shall be presented in Section IV to demonstrate the use of the models developed by this study.

B. DATA NORMALIZATION

Once the purpose has been determined, the cost estimation process moves into its second phase, *data normalization*. Typically a cost model predicts costs of new systems based on underlying relationships discovered from historical systems. These relationships will form the basis of the eventual model. They must be grounded in reality

by normalizing the data or the cost estimate will not be credible. In particular, the historical data must be normalized for *content, quantity* and *inflation.*

1. Normalization for Content

Before any other adjustments are made, the data must be verified to be approximately comparable to one another, both physically and programmatically. This is the logical "apples-to-apples" argument, ensuring that each item in the data set is a member of the same underlying population as every other item in the set. As an example, a database that includes the procurement costs of frigates and the reactivation cost of battleships would not be appropriate for predicting the procurement cost of a new destroyer, despite the physical similarities of each historical element. The battleship costs include only the upgrade costs for an existing ship, while the frigate costs include the production of an entirely new ship. This does not imply that the data cannot reflect differences among the included systems, but rather that at some functional level they must be equivalent. The comparison is typically made using a work breakdown structure (WBS), which is an outline of program costs partitioned into various hierarchical subcategories. A WBS for Naval Ships is shown in Table 1. (MIL-HDBK-881) Cost analyses may address the WBS at almost any level, from detailed divisions of subcategories of program elements to broad overviews summarizing entire programs.

Ship System Work Breakdown Structure		
Level 1	**Level 2**	**Level 3**
Ship System		
	Ship	
		Hull Structure
		Propulsion Plant
		Electric Plant
		Command and Surveillance
		Auxiliary Systems
		Outfit and Furnishings
		Armament
		Integration/Engineering
		Ship Assembly and Support Services
	Systems Engineering/Program Management	
	System Test and Evaluation	
		Development Test and Evaluation
		Operational Test and Evaluation
		Mock-ups
		Test and Evaluation Support
		Test Facilities
	Training	
		Equipment
		Services
		Facilities
	Data	
		Technical Publications
		Engineering Data
		Management Data
		Support Data
		Data Depository
	Peculiar Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Common Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Operational/Site Activation	
		System Assembly, Installation and Checkout on Site
		Contractor Technical Support
		Site Construction
		Site/Ship/Vehicle Conversion
	Industrial Facilities	
		Construction/Conversion/Expansion
		Equipment Acquisition or Modernization
		Maintenance (Industrial Facilities)
	Initial Spares and Repair Parts	

Table 1. Example Work Breakdown Structure.

2. Normalization for Quantity

When analyzing cost data from several different systems, the associated production quantities play a significant role. Each new unit coming off a production line typically costs less than the units produced before, as workers and supervisors learn from experience and improve in efficiency. In learning curve theory, the production cost of a unit is reduced by a constant percentage each time the production quantity is doubled. In order to compensate for this effect, it is desirable to relate all costs to a common point of production, such as the theoretical first unit cost (T1), when comparing production cost data. The T1 cost may differ from the actual cost of the first-produced unit. An example of actual costs, the fitted curve and the resulting T1 is shown in Figure 2. (USALMC, p. 7-1 to 7-3)

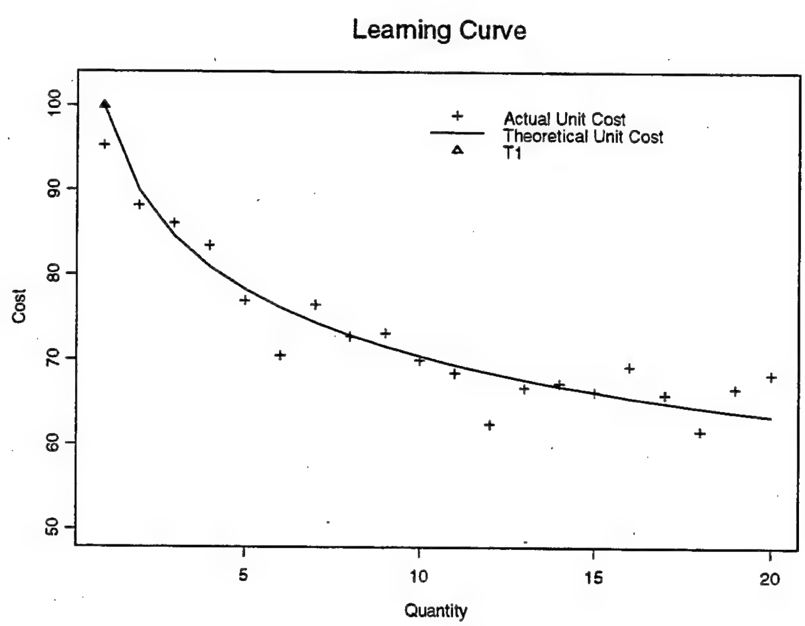


Figure 2: Example Learning Curve.
Showing actual and theoretical costs.

3. Normalization for Inflation

Because any particular monetary unit 'doesn't buy what it used to,' costs from different programs need to be adjusted to a common time reference to be compared. Often, historical data describe how money was actually spent during a given year of a program. The dollar values are current dollars, and reflect the purchasing power of a specific amount of money in each given year. For the example program in Table 2, the nominal costs of the two programs are equal. However, the Program 1 dollars were spent ten years before those in Program 2. During those ten years, inflation has reduced the value of a dollar in purchasing a product. Thus Program 2 has purchased less, with its less valuable dollars, than Program 1.

	Program Year Spending ($)					
Program #	1970	1971	1980	1981	1982	Total
Program 1	25	75				100
Program 2			30	30	40	100

Table 2. Unadjusted Program Spending.
Showing spending in budget year dollars.

The solution is to adjust each yearly total to a common time reference. This adjustment is made using tables designed to convert between different years, using historical cost changes in specific economic commodities to measure the change in value of a dollar over time. (USALMC, p. 11-1) These adjusted values are called constant year (CY) values, and represent the price of acquiring a particular product in a specific year. Because labor wages and other factors change at different rates for different products, the tables are tabulated for particular Naval program areas, such as ship construction (SCN), weapons construction (WPN) or aviation programs (APN). A properly adjusted

comparison, assuming SCN program dollars, is shown in Table 3, clearly showing that Program 1 is a more expensive program than Program 2.

	Constant Year Spending (CY98$)					
Program #	1970	1971	1980	1981	1982	Total
Program 1	149.25	410.73				559.98
Program 2			63.83	57.20	66.72	187.75

Table 3. Adjusted Program Spending.
Showing spending in constant year dollars.

C. DATA ANALYSIS

Once the purpose of the cost estimate has been chosen and the data normalized, data analysis can be used to identify the relationships between the historical cost data and the specific attributes of the historical systems.

1. Variable Selection

The first task is to select variables suitable for predicting the cost. Parametric methods are often viewed as a panacea for understanding the reason behind a particular effect. Unfortunately, the view is often misguided—the relationships demonstrated by parametric analysis establish associations, but not causality. As an example, modern grenades have become both smaller and more lethal. However, reducing the size of munitions will not increase their lethality. A separate factor, technological improvement, accounts for the trend between size and lethality.

The independent variables chosen for parametric analysis should *cause* the changes in the dependent variable. When demonstrating that a relationship is causal instead of associative, three concepts must be addressed. First, the relationship between the independent and dependent variables must be consistent; when other things are equal in a population, the relationship should consistently differ in a specific direction, or even

in magnitude, when the independent variable is adjusted. For instance, holding all characteristics of a product constant except the number purchased, buying two items usually costs more, often exactly twice as much, as buying one. Second, the relation must be responsive to changes in the independent variable; altering the independent variable should cause a change in the dependent variable. Doubling the weight of a satellite should increase the cost of getting it into orbit. Finally, a mechanism, obvious or not, should be responsible for the change. (Mosteller & Tukey, p.260-1) For example, antennas designed for higher frequencies are smaller than ones designed for low frequencies, since a relationship exists between the surface area of an antenna and its frequency. Conversely, although a platform may have fewer large search radar antennas than small fire-control antennas, the relationship between size and number is only an association—the number of antennas is determined by mission need.

The assurance of causality is best established not by statistics, but by expert opinion. While not infallible, experts often have the practical experience necessary to separate the prospective causal factors from the myriad associative ones. They might also offer guidance as to the mathematical form such relationships take; certain parameters vary linearly; others vary exponentially, requiring transformation to coax them into a linear form suitable for regression analysis.

In addition to the causal nature of the independent variables, the ranges over which the historical observations occur must also be considered. A regression model calculates the line which best fits the points in the data set. Extrapolating this relationship outside the range of the data extends it into new areas where the relationship

may no longer hold. Most physical phenomena are subject to this problem. In electrical theory, for example, a resistor develops a voltage across it directly in proportion to the amount of current that passes through it. The relationship may be verified by varying the current and plotting the developed voltage. But driving an exceptionally large current through a small resistor will not develop a proportional voltage across the resistor—it will simply turn the resistor into smoke and gas. To predict a relationship outside the range of the data raises serious questions of credibility and should be avoided whenever possible.

2. Relationship Determination and Transformation

The importance of seeking a linear relationship cannot be overstated. When regression techniques are applied to a data set, they will identify all trends as linear functions. If a variable shares a non-linear relationship with the dependent variable, it must be transformed to make the relationship linear or the regression will exhibit excessive error. A pair-wise examination of the independent variables against the dependent variable may show evidence for or against the claim that a relationship between them is linear. Caution should be exercised however; transformations make the model difficult to interpret—log(hours) are not an intuitive measure of time. Also, with relatively small data sets, the determination that a relationship between two variables is non-linear is subjective at best.

3. Regression Model Postulation

After collecting, normalizing and transforming the data, statistical analyses may be employed to identify the underlying relationships which show promise as predictors

for the dependent variable. Foremost in this category is ordinary least-squares (OLS) regression. OLS reduces a collection of points into a set of coefficients defining a line. With OLS regression, the sum of squared vertical distances from each data point to the line is made to be as small as possible. Each regression forms a mathematical model, which takes as inputs the independent variables of the OLS regression and returns an estimate of the dependent variable. The number of independent variables used in a regression formulation provides a convenient classification scheme.

a. Single Variable Models

Single variable models are the simplest linear regressions. They describe the dependent variable, often cost, as a linear function of a single independent variable. Because they may be fully described in two dimensions, they are easy to display and explain. Also, they provide invaluable insight during high-level studies when detailed information about new systems, required to fulfill the data requirements of a multivariate model, cannot be reasonably generated. However, if several independent variables are available, the additional information that may be contained in the remaining independent variables is lost. An easy solution to the problem of lost information would be to include more variables. However, this solution creates new problems of its own, as will be described in the next paragraph.

b. Multiple Variable Models

Multiple independent variable models often describe the dependent variable better than single variable models. The additional information enables the multivariate model to make predictions of the dependent variable with greater accuracy

than a univariate model. However, the multiple variables often interact behind the scenes to mask their effectiveness in predicting the dependent variable. Variables are considered correlated if a relationship exists between them, i.e. if knowing some information about one variable offers some information about the other. An example shows the possible errors associated with correlation (ρ) in a regression model. Consider a study where cost is being predicted by two variables, weight and length. In this example, the correlation between the two variables is 1.0, indicating that by knowing one variable, the other one is completely known as well (weight is directly proportional to length). Several univariate mathematical relationships describe the models of Figure 3:

1) Cost = 1*Weight
2) Cost = 1*Length
3) Length = 1*Weight

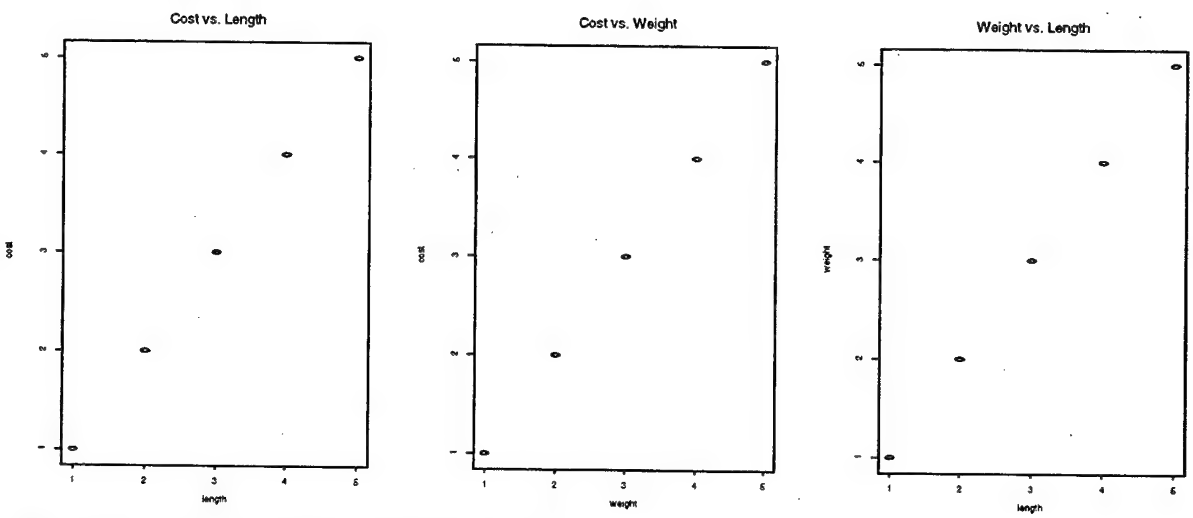


Figure 3: Correlated Multivariate Data.
Showing relationships between dependent and independent variables.

Because of the collinearity between length and weight, several multivariate models perform equally well. The models themselves however appear contradictory.

Model 1: Cost = 1.5*Weight – 0.5*Length
Model 2: Cost = -0.5*Weight + 1.5*Length
Model 3: Cost = 0.5*Weight + 0.5*Length

All describe the relationship perfectly. However, the models should be used only if the collinear relationship between the two independent variables (the direct proportionality between weight and length) holds for the new data. If not, the model predictions are suspect. Note for the above models, an object with weight=1 and length=3 (the collinear relationship is violated) would cost zero, four or two dollars, depending on the model used to predict a cost. This does not immediately disqualify a model with a high ρ, but cautions that the relationship between the correlated variables must also be found within the new data before that new data may be used with such a model for predictive purposes.

D. MODEL DETERMINATION AND CER SELECTION

The goal of every regression strategy is to produce a simple expression relating cost as the dependent variable to one or more independent variables. Although it would seem that regression analysis should fulfill this objective easily, the regression techniques must be justified by further analysis.

The statistics generated by OLS regression will be used to justify both the *form* and the *coefficients* of the model. The form of a model consists of the independent variables used to make the model. When justifying a model form, the analyst will decide which variables to include, as well as the appearance they should take—whether sums, products, ratios or other combinations. In addition, the analyst must verify the actual

coefficients of each independent variable. The cost estimating relationship (CER) includes both a form and specific values for all coefficients.

1. Justifying the Model Form

Determining a model form consists of deciding which of the independent variables should be included in the model. It is assumed that the variables were selected because they have a causal relationship with the dependent variable that would help predict new values. With the data in hand, that assumption may be tested. The regression returns statistics to back up or refute the expected relationship. These statistics provides an indication about whether particular variables should be included in a model, enabling the analyst to sort, build and shrink regression models by adding, removing and combining variables until an acceptable form is found.

Two indicators of the acceptability of a model form are the p-values associated with the F statistic and the p-values associated with particular variables, or t statistics. The p-value may be interpreted as follows: when a gambler asserts that three rolls of a die will result in three sixes, one might consider such an event to be unlikely or rare, assuming the die is fair. Observing him roll a six on all three attempts raises the question of whether the die is indeed fair. The chances of having a six occur three times in a row are $(1/6)^3$, or 0.0046. An event this rare or more so should only happen about once in every 200 tries. This backs up the suspicion that the die is probably not fair. So, a p-value may be interpreted as the probability of seeing an event this rare or more so if the assertion being made is true. The significance of the p-value is measured by its magnitude. In the example the significance of the fairness of the die is 0.0046.

The p-value associated with the F statistic can be interpreted as the probability that the coefficients of the independent variables in the model are all zero. In such a case, the average cost ($\bar{y}$) would provide an equally accurate estimate of a predicted cost. The p-value for the t statistic associated with each independent variable describes a similar probability, but with respect only to the coefficient of a particular independent variable.

Multivariate models cannot rely only on the F statistic to determine whether a model is acceptable. One common strategy for generating prospective multivariate model forms is backward elimination. Backward elimination first generates a composite model by performing a regression of the dependent variable against all independent variables, then systematically eliminates individual independent variables until a model is found in which all coefficients are reasonably significant.

The level of significance at which remaining variables are deemed worth keeping, or the required significance (α), is a subjective determination. Although an α level of 0.05 is a common requirement in scientific analysis, this study shall select a less restrictive level of 0.2 as a maximum acceptable variable p-value significance. The reason for increasing the required α to 0.2 is two-fold. First, the required accuracy of a cost model does not usually require a 5% tolerance. The goal is not to eliminate all variables that do not explain a majority of the change of the dependent variable, but to identify all variables that appear to contain information that assists in the prediction of the dependent variable.

Second, the meaning of the p-value has been skewed because of the method used to generate the models. When a regression is conducted once, the p-value describes the

chances of seeing a result as rare or more so, given that a model describes the data just as well without the variable in question. With an α=0.05, the rare event happens only once in twenty times. But when regressions are conducted repeatedly, the odds of the rare event increase significantly. If twenty regressions are performed, it would be unlikely for the rare event not to occur. Thus the actual significances are larger than the p-values reported by the statistical tests. Arbitrarily setting a low α will not assure model parameters are significant, only that additional parameters shall be eliminated.

The larger α poses no serious problem. If an insignificant variable is accidentally included in the model, the true coefficient associated with the variable would be zero. Including such a variable does not change the prediction. On the other hand, if a significant variable is omitted, the model becomes biased; a change to the omitted variable causes the dependent variable to change and the model's prediction will be in error, not just by chance, but specifically because of the changing omitted variable (Hamilton, p. 73).

If a model contains only variables whose individual p-values are <0.2, the overall model F-statistic p-value will also be smaller than 0.2, as the probability of several unlikely events happening simultaneously is always smaller than the probability of any of the events individually.

2. Justifying the Coefficients of the Model

In addition to justifying the model *form*, the *coefficients* of the model, as well as their signs, must be considered when deciding whether a particular model is acceptable.

Two areas must be evaluated: the assumptions of the model and how well the model fits the data it was built around.

a. Evaluating the Assumptions of the Model

A linear regression describes the value a dependent variable should take given a particular set of independent variable values. OLS regression supplies the best way to describe a data set with a linear model, provided certain conditions are met. If the conditions do not hold, the results of OLS become less trustworthy. In these cases, OLS may still provide insight into a database, but might not provide the best description of the data. As the assumptions are disobeyed, the OLS model becomes progressively less capable of describing the data. (Hamilton, p. 109)

Under OLS, every dependent variable may be written as a linear combination of the independent variables, together with a random error term. The error term explains all variations in the dependent variable not caused by the independent variables in the equation, and is often named the residual. Equation 1 summarizes this relationship.

$$Y_i = \beta_0 + \beta_1 X_{(i,1)} + \beta_2 X_{(i,2)} + \ldots + \beta_k X_{(i,k)} + \varepsilon_i \qquad [1]$$

Y_i : actual dependent variable data value
β_j : coefficient for independent variable j
$X_{(i,j)}$: i^{th} individual independent variable data value for variable j
ε_i : i^{th} error term or residual

The linear regression depends upon the validity of five underlying assumptions. They are:

1) Every variable that causes the dependent variable to change is in the model. Because of this, a given set of independent variables shall always produce the same result, along with an error term.
2) The error terms will have a mean of zero.
$E[\varepsilon_i] = 0 \quad \forall\, i.$
3) The error terms will have constant variance.
$Var[\varepsilon_i] = \sigma^2 \quad \forall\, i.$
4) Error terms are not correlated with one another.
$Cov[\varepsilon_i, \varepsilon_j] = 0 \quad \forall\, i \neq j.$
5) Error terms are normally distributed. (with a mean of zero and a constant variance)
$\varepsilon_I \sim Normal(0,\sigma^2) \ \forall\, i.$

(Hamilton, p.110-3)

Unfortunately, with sample data, two of these assumptions cannot be verified (Hamilton, p. 112-3). Assumption (1) assumes perfect knowledge about the relationship between the dependent and independent variables—this kind of assurance can never be provided by science, regardless of the topic or application. Similarly, Assumption (2) can never be verified in practice—if the error terms have a non-zero mean μ', all predictions would miss the true dependent variable value by μ'. However, when this situation is estimated, that discrepancy would be corrected by modifying β_0. The following two situations are indistinguishable: $\{E[\varepsilon_i] = 0$ with $\beta_0 = c\}$ and $\{E[\varepsilon_i] = \mu' \neq 0$ with $\beta_0 = c - \mu'\}$.

The remaining three assumptions should be investigated using analytic techniques and diagnostic plots. If Assumption (3) does not hold, a condition known as heteroscedasticity, the variance of the model will be estimated overly high or low, making estimates of confidence intervals inaccurate. (Hamilton, p. 113) Heteroscedasticity among the ε_i may be seen easily in a plot of predicted dependent

variables against $|\varepsilon_i|$, when the average residual magnitude is not constant over the range of predicted dependent variable values.

If the data violates Assumption (4), and error terms are correlated, the model's variance will also be affected. The difficulties of a correlated model have already been described when discussing multivariate models. Correlation is best detected using a covariance matrix of the independent variables. Real data can be expected to show some correlation; a $|\rho|<0.3$ should not be a concern. If $|\rho|>0.7$, the correlation must be addressed. To put the effects in perspective, if two variables in a model are mildly correlated ($\rho=0.3$), the actual standard error could be 105% of the reported standard error. If $\rho=0.7$, the actual standard error could be almost 140% of the reported value. (Hamilton, p. 113, 133-6)

In a similar way, violations of Assumption (5), or non-normality, also make the model less accurate—calling into question the p-values of both the t and F statistics. (Hamilton, p. 112-3) Since the residuals are supposed to have a normal distribution, with a particular mean and variance, any non-normality may be detected using a quantile plot of the residuals. The plot compares the fraction of residuals that are smaller than each quantile of the normal distribution with the same mean and variance. A straight line on the quantile plot indicates the residuals are indeed normal.

b. Evaluating the Fit of the Model

Three additional statistics, the coefficient of determination (R^2), the residual standard error (*RSE*) and the coefficient of variation (*CV*), offer insight into how well a model fits the data around which it was built. Each presents similar information in

a different way. The *RSE* provides a measure of the typical deviation of an actual data point from the predicted value on the regression line. The calculation returning the *RSE* is similar to the calculation of a variance or standard deviation. Equation 2 describes the process:

$$RSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n-k-1}} \qquad [2]$$

RSE : Residual Standard Error
n : number of data points in database
k : number of independent variables
y_i : actual dependent variable value
$\hat{y}_i$: predicted dependent variable value

The *RSE* may be used as an estimate for the standard error of a predicted value from a model, and is useful in calculating uncertainty and confidence interval information about model predictions. (Hamilton, p. 36)

The coefficient of variation (*CV*) places the *RSE* in perspective by describing the ratio of the *RSE* to the average value of the dependent variable. Less formal than a confidence interval, the *CV* returns the relative size of the error to the average value of the dependent variable. The *CV* thus describes the expected percentage error of a prediction. The coefficient of variation may be determined using equation 3.

$$CV = \frac{RSE}{\bar{y}} \qquad [3]$$

CV : coefficient of variation of model
RSE : residual standard error of model
$\bar{y}$: average value of dependent variables

The coefficient of determination (R^2) is a ratio of the explained variation to the total variation. R^2 values range from zero, for a model that explains none of the variation shown by the dependent variable, to 1.0, for a model that completely describes the data used to generate the model. The coefficient of determination may be calculated using Equation 4.

$$R^2 = \frac{\Sigma(\hat{y}_i - \bar{y})^2}{\Sigma(y_i - \bar{y})^2} = 1 - \frac{\Sigma(\hat{y}_i - y_i)^2}{\Sigma(y_i - \bar{y})^2} \qquad [4]$$

y_i : actual dependent variable value
$\hat{y}_i$: predicted dependent variable value
$\bar{y}$: average value of dependent variables

The coefficient of determination will always increase as additional independent variables are added. The new variables cannot make the fit of the model worse, but they provide some additional information about the makeup of the data set. However, adding variables simply to raise R^2 does not necessarily improve the model. An extreme case would add a binary variable for every data point in the database. Such a model would have an R^2 of 1.0, as it perfectly describes the data set, but would not offer any information about a new data point that did not correspond exactly to a previous point. Accordingly, to balance the improvement in R^2 to the cost of using additional parameters, the coefficient of determination can be adjusted. The adjusted R^2 accounts for the model size relative to the sample size by reducing the R^2 by a fraction of the

unexplained model variation, and is described in Equation 5. In general:

$$R_a^2 = R^2 - \left(\frac{k-1}{n-k}\right)(1-R^2) \qquad [5]$$

R_a^2	: adjusted coefficient of determination
R^2	: coefficient of determination
k	: number of model parameters, including intercept
n	: sample size

Note as k approaches n, R_a^2 is reduced towards zero and can even become negative. This adjustment emphasizes the objective of the model, predicting future values instead of simply describing the current data. Any model can describe the database by using the database—only by identifying the underlying relationships with CERs may a model be applied effectively to new data.

E. MODEL SENSITIVITY AND CROSS-VALIDATION

Once a model has been created that justifiably describes the data set, the question of how well it predicts new values may be addressed. The statistics from the model justification are often used to show that it will work in a new setting with new data. But those statistics actually describe only how closely the model fits the old data used in its construction. OLS will make the best use of the information contained in the data set—both the underlying relationships between the variables and the patterns that occur simply by chance.

Consider the database; from a statistical perspective, the points have been selected from an infinite number of possible choices. Each point consists both of information explained by the model and random error. However, the error terms will, through random chance, form patterns periodically, as though they were information. The OLS

procedure does not waste any of this information—it will describe the patterns within the data set, whether the cause of the pattern was real or happenstance. Therefore, the model will better describe the data from which it was built than any new data used later. When the sample size is relatively small, the chances of improperly characterizing a relationship are greater still. In order to gain some appreciation of a model's ability to predict new data, it must be validated. Validating a model by evaluating its ability to predict new data is called cross-validation. (Mosteller & Tukey, p. 36-7)

1. Cross-validation Performance

Cross-validation may be performed in several ways, under the categories of 'simple' or 'double' cross-validation. The categories reflect the degree to which the new data has been previously studied or used.

a. Double Cross-validation

The fundamental way to perform cross-validation, often described as double cross-validation, involves acquiring new data after the form and coefficients of the model have been determined. Alternatively, the data set can be separated and a portion withheld before any examination has taken place. The new data is held in reserve until the models are completely determined, then the withheld data are entered in the model and the predicted values are compared with the actual dependent variable values. The difference between the actual and predicted values represents the performance of the models on entirely new data. (Mosteller & Tukey, p. 36-38)

b. Simple Cross-validation

Unfortunately, many data sets are too small to be cut into pieces and still produce worthwhile models. Recalling that adjusted R^2 is a function of the relative size of model to the database, using a subset may eliminate any hope of statistical significance. In such a case, simple cross-validation provides a reasonable alternative. In simple cross-validation, the data are partitioned into several (r) subsets of approximately equal size (n') after the determination of a model's form has been made. Withholding one of the subsets, the particular coefficients for the model are determined from the remainder. The model is then used to predict the dependent variables in the withheld subset, using the subset's associated independent variables. Each prediction will miss by a particular amount, from which a squared residual may be calculated. The squared residual is calculated as in the *RSE*, the square of the difference between the actual and predicted value of the dependent variable. This process is repeated in turn with each subset, noting the squared residual of each. The average squared residual provides some measure of the quality of a regression on a data set of size $n - n'$. Because regression performs better on large data sets, this process can be maximized by withholding only a single data point, creating a model with the remaining data, predicting the excluded point and repeating the process until every point has been excluded (Mosteller & Tukey, p. 38-9). The average of the squared residuals may be used to calculate a cross-validated *RSE* that approximates the expected performance of the model

when predicting new data. This cross-validated *RSE* may also be used in the equation for *CV*. The equation for cross-validated *RSE* (RSE_{cv}) follows as Equation 6.

$$RSE_{cv} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i'\right)^2}{n-k-1}} \qquad [6]$$

RSE_{cv} : cross-validated Residual Standard Error
n : number of data points in database
k : number of independent variables
y_i : actual dependent variable value
$\hat{y}_i'$: predicted dependent variable value using subset model

The sum of squared residuals may be used as in the Equation 3 to calculate a cross-validated coefficient of determination (R_c^2). The equation for a cross-validated coefficient of determination is Equation 7.

$$R_c^2 = \frac{\sum\left(\hat{y}_i' - \bar{y}\right)^2}{\sum\left(y_i - \bar{y}\right)^2} = 1 - \frac{\sum\left(\hat{y}_i' - y_i\right)^2}{\sum\left(y_i - \bar{y}\right)^2} \qquad [7]$$

R_c^2 : cross-validated coefficient of determination
y_i : actual dependent variable value
$\hat{y}_i'$: predicted dependent variable value using subset model
$\bar{y}$: average value of dependent variables

The cross-validated coefficient of determination may also be adjusted using Equation [5]. Although simple cross-validation does not actually demonstrate the model's future performance, it provides a better measure of the predictive qualities of the models than *RSE* and R_a^2 alone.

THIS PAGE INTENTIONALLY LEFT BLANK

III. METHODS

The methodology outlined in Section II must now be tailored to create predictive cost models that turn estimated data into quantitative predictions. Section III details the specific procedures and decisions required when generating a cost estimation model that will accept such approximate data.

A. PURPOSE DEFINITION AND DATA COLLECTION

The purpose of the cost estimate will determine most aspects of the study. But the availability of data will drive the ability to create useful models. The cost data used in this study was the most comprehensive available at the time.

1. Defining the Purpose

This study shall focus on the creation of a parametric cost estimation model that converts approximate and uncertain estimates about Naval ship parameters into an average ship procurement cost estimate, including a measure of uncertainty. Cost will be the dependent variable, with physical and performance parameters serving as independent variables. The models are intended to be used in force structure cost analysis, particularly in the context of new force activation and acquisition; and reorganization, modification and modernization, as defined in Section II under Cost Analysis Applications. As such, the information about ship parameters is expected to be rough and incomplete; however, the cost models should be able to generate answers even in the face of limited information. The models will also be able to describe the expected variability of their estimates.

2. Data Collection

Because the parametric model will be expected to generate cost estimates for future systems that have not yet been designed, the model must be built from general data. While identifying the particular relationships specific to guided missile destroyers (DDGs) would provide additional accuracy when predicting the cost of new DDGs, the models for this study must be able to predict a wide variety of platform types. Therefore, the data includes as many classes and spans as much historical ground as possible.

a. Cost Data

The reference, U.S. Weapon System Costs (Data Search Associates), tabulates procurement cost data for several ship classes. The entries represent major Naval ship acquisitions from 1973 to the present. All shipbuilding programs from the aforementioned tables have been incorporated into the cost data set. Several of the entries had missing ship class names and dated or inconsistent ship class designators, but the errors were easily corrected using supplied shipbuilder or contractor information and program start year data. The ship classes included in the data set are summarized in Table 4.

Class	Class Name	Data Duration
AD41	Yellowstone	1975-1981
TAE26	Kilauea	1995-1997
AGOR23	Thomas G. Thompson	1985-1999
TAGS60	Pathfinder	1995-1996
AO177	Cimarron	1976-1983
AOE6	Supply	1985-1993
ARS50	Safeguard	1979-1985
CG47	Ticonderoga	1976-1988
CVN68	Nimitz	1973-1999
DDG51	Arleigh Burke	1981-1999
DD963	Spruance	1973-1983
FFG7	Oliver Hazard Perry	1975-1986
LCACł	not applicable	1982-1994
LHD1	Wasp	1981-1999
LPD17	San Antonio	1996-1999
LSD41	Whidbey Island	1987-1993
MCM1	Avenger	1979-1991
MHC51	Osprey	1985-1993
MKV	not applicable	1994-1998
MSH1	Cardinal	1982-1984
PHM1	Pegasus	1973-1978
SSBN726	Ohio	1973-1991
SSN688	Los Angeles	1973-1991
SSN774	Virginia	1992-1999
SSN21	Seawolf	1985-1999
TAGOS1	Stalwart	1978-1990
TAO187	Henry J. Kaiser	1980-1991
TAFS	Lyness	1982-1984
TARC7	Zeus	1979-1983
TATF166	Powhatan	1975-1979

Table 4. Ship Classes.
Including class name and procurement period.

b. Weapon System Class Parameters

Performance and technical parameters of the naval ship classes have been obtained from JANE's Fighting Ships (JANE's Publishing) and verified using Ships and Aircraft of the U.S. Fleet (Polmar). Seven attributes were chosen for inclusion in the data set:

1) Length (*LEN*), the overall length of the craft in feet.
2) Light Displacement (*DISP*), the weight in tons of the ship hull, machinery, equipment and spares. (Transportation Institute)
3) Beam (*BEAM*), the vessel width at its widest point, in feet.
4) Number of Engines (*ENGNUM*), the number of engines used for propulsion.
5) Propulsion Type (*PROP*), the engine type used for propulsion. d: diesel, s: steam, t: gas turbine, n: nuclear power.
6) Shaft Horsepower (*SHP*), the total engine power, in hp, used for locomotion.
7) Maximum Speed (*MAX*), the published maximum speed of the craft in knots.

These attributes represent the general information that might be available or estimable for a Naval ship long in advance of specific designs. Detailed information cannot always be expected when performing a force structure cost analysis. General parameters allow the analyst to identify aspects of force structure elements that may be similar to historical craft in the database. The data for each class, together with average cost given in constant 1999 dollars (CY99M$), is shown in Table 5.

Ship Class	*AVGCOST* (CY99M$)	*NUM*	*DISP* (tons)	*LEN* (feet)	*BEAM* (feet)	*ENGNUM*	*PROP*	*SHP* (khp)	*MAX* (knots)
AD41	852.265	4	13318	643.8	85	2	s	20	20
TAE26	35.354	3	9238	563.8	81	3	s	20	22
AGOR23	131.306	7	2100	274	52	3	d	6	15
TAGS60	24.417	3	3019	329	58	2	d	8	16
AO177	349.113	5	8210	708.3	83	1	s	24	19.4
AOE6	559.479	4	19700	754.8	107	4	t	100	26
ARS50	124.418	4	2725	254.9	51	4	d	4.2	13.5
CG47	1459.251	27	7015	567	55	4	t	86	30+
CVN68	5107.259	7	77606	1092	134	4	n	260	30+
DDG51	971.560	47	6682	504	66.9	4	t	105	32
DD963	606.673	15	6156	563.2	55.1	4	t	86	33
FFG7	484.068	50	2770	445	45	2	t	41	29
LCAC1	34.089	84	102.2	81	43.7	4	t	15.82	50
LHD1	1405.056	7	28233	844	106	4	s	77	24
LPD17	810.509	4	25300	683.7	104.7	4	d	40	22
LSD41	330.789	4	11125	609.6	84	2	d	41.6	22
MCM1	154.282	17	1195	224.3	38.9	2	d	2.6	12.5
MHC51	180.747	9	803	187.8	35.9	2	d	1.16	12
MKV	8.626	18	75	82	18	2	d	4.506	50
PHM1	228.469	4	198	131.5	28.2	1	t	16.77	50
SSBN726	2362.590	19	16600	560	42	2	n	60	24
SSN688	944.541	51	6082	362	33	2	n	35	32
SSN774	3475.525	2	7700	377	34	2	n	24	28
SSN21	2212.306	3	7460	353	42.3	2	n	52	35
TAGOS1	82.593	23	1600	224	43	4	d	3.2	11
TAO187	211.402	18	9500	677.5	97.5	2	d	32.54	20
TARC7	473.616	1	8297	502.5	73.2	5	d	12.5	15.8
TATF166	47.389	7	2000	240.5	42	2	d	4.5	15

Table 5. Ship Class Physical and Performance Parameters.
Showing all prospective dependent and independent variables.

B. DATA NORMALIZATION

Because the data includes a variety of ship classes, proper normalization is of key importance. Each data point must be carefully evaluated to ensure it is equivalent to every other in terms of *content*, *quantity* and *inflation*.

1. Content Normalization

Cost data may be normalized using the WBS of the platform in question. The Weapon System Cost data (Data Search Associates) divides procurement costs into two categories: *Procurement Costs* and *Other Procurement Costs*.

Procurement Costs include costs of all WBS Level 2 categories from Table 1. except Training, Peculiar Support Equipment and Common Support Equipment. It also includes all costs, both contract and in-house of the Production Non-recurring and Recurring cost categories, including allowances for engineering changes, warranties and first destination transportation, unless the latter is a separate budget line item.

Other Procurement Costs represent the costs of outfitting the ships. The costs include spares, repair parts, escalation and cost growth, post-delivery and other material required for storeroom and operating space initial allowances. It also includes design, planning, government-furnished materials and related labor costs required to correct sea-trials deficiencies.

The performance and technical data were also investigated to ensure that measurements from one ship class corresponded to similar measurements from another. Overall length was chosen instead of waterline length because it does not depend on ship draft. Beam measurements, as used, do not include protrusions such as the flight deck or bridge wings. The engine number consists of the steam or gas turbines used for locomotion and diesel engines used directly or indirectly for propulsion. The maximum speed figures are unclassified estimates. Several classes did not list a particular top speed, listing instead 30+, indicating an unspecified speed in excess of thirty knots.

Substituting thirty knots would not properly reflect the actual capabilities of the classes, and inputting an N/A effectively removes those classes from consideration at all with respect to the *MAX* independent variable. As a compromise, a figure was obtained by using the highest published speed of a similar craft, the DD 963. Because of this, a speed of 33 knots was used for the CVN 68 and CG 47. Although imprecise, the estimate provides some accuracy without making the data classified.

Two programs were found to be incompatible with the remaining data and were stricken, as they did not represent actual new-production programs. The MSH-1 program was cancelled before entering production, and the TAFS program dollars were used only to convert existing vessels rather than to produce new ones.

Additionally, five classes were deemed too dissimilar from the rest of the data to be included. The nuclear powered vessels—four submarines and one carrier class do not obey the same production cost rules as conventional ships. They differ from remaining ships both in value and trend, reflecting their unique production methods, quality assurance requirements, labor costs, environmental support and other factors.

2. Quantity Normalization

While normalization of cost figures to theoretical first unit cost (*T1*) values for comparative purposes provides the most accurate representation of skilled production on unit cost, the process requires individual procurement costs for individual vessels from individual shipyards. Unfortunately, the cost data available for this study does not include such information. The data is tabulated in a 'by-year' format and cannot be separated into specific units, lots or shipyards.

While not ideal, the data are sufficient to create an estimate for hypothetical alternatives and components of strategic plans. If the objective of the study were to predict the cost of producing an additional number of ships from a well-defined program, the use of averaged values might be questionable. However, the objective is a model that predicts cost from uncertain inputs. In such circumstances, inflation-adjusted averaged cost figures provide an acceptable compromise between cost accuracy and data availability.

3. Inflation Normalization

The tabulated cost figures consist of actual dollars spent in a given year. To normalize all values for inflation, individual amounts are converted into constant 1999 (CY99) dollars using "Inflation Indices and Outlay Profile Factors" prepared by the Naval Center for Cost Analysis (NCCA). These normalized figures are then summed by class and divided by the number of craft produced in the program, yielding an average procurement cost in CY99M$. Values for the average procurement costs of all classes are also included as the variable *AVGCOST* in Table 5.

As an example, the Kilauea class (TAE 26) data covers the years 1995 to 1997. The procurement cost from each year is adjusted to CY99 by multiplying by the inflation index for the particular year. The adjusted dollars are then added and divided by the total quantity of ships produced within that time frame, resulting in an inflation adjusted average cost for the TAE 26 class. The data is summarized in Table 6.

Procurement Year	Procurement Cost (BYM$)	Inflation Index	Adjusted Procurement Cost (CY99M$)	Quantity Produced
1995	30.3	1.0899	32.70	1
1996	30.0	1.0685	32.38	1
1997	39.2	1.0455	40.98	1
	Total Cost		106.06	
	Total Quantity		3	
	Average Cost (CY99M$)		35.35	

Table 6. Inflation Adjustment for TAE 26 Class.
Converting yearly budget spending into an adjusted class average.

C. DATA ANALYSIS

The data analysis follows the procedure outlined in Section II C. Deviations from the outline shall be explicitly described and justified.

1. Relationship Determination and Transformation

Linear regression best explains the relationships between independent and dependent variables if the relationships are linear. If the relationships are not linear, they must be transformed or the models will exhibit excessive error. No preconceived inferences are made as to the expected relationship form. The statistical package S-Plus provides a function called loess that may be used to subjectively evaluate the linearity of the data. Loess is a locally weighted regression that picks the best line segment to describe only the points in the immediate vicinity of every point. (S-Plus Guide to Statistics, p. 159-60) When plotting both the OLS line and loess line, departures between

the two indicate local behavior that does not match the overall linear trend of the OLS line. If the loess line significantly strays from the OLS line, transformations may be required to convert the data into following a linear relationship. Both the OLS line and loess line for conventional ships are shown in Figures 4 and 5. Although the loess lines do not mirror the OLS lines exactly, they do not exhibit any particular behavior to suggest a non-linear relationship. Therefore, all variables will be used in models without transformation.

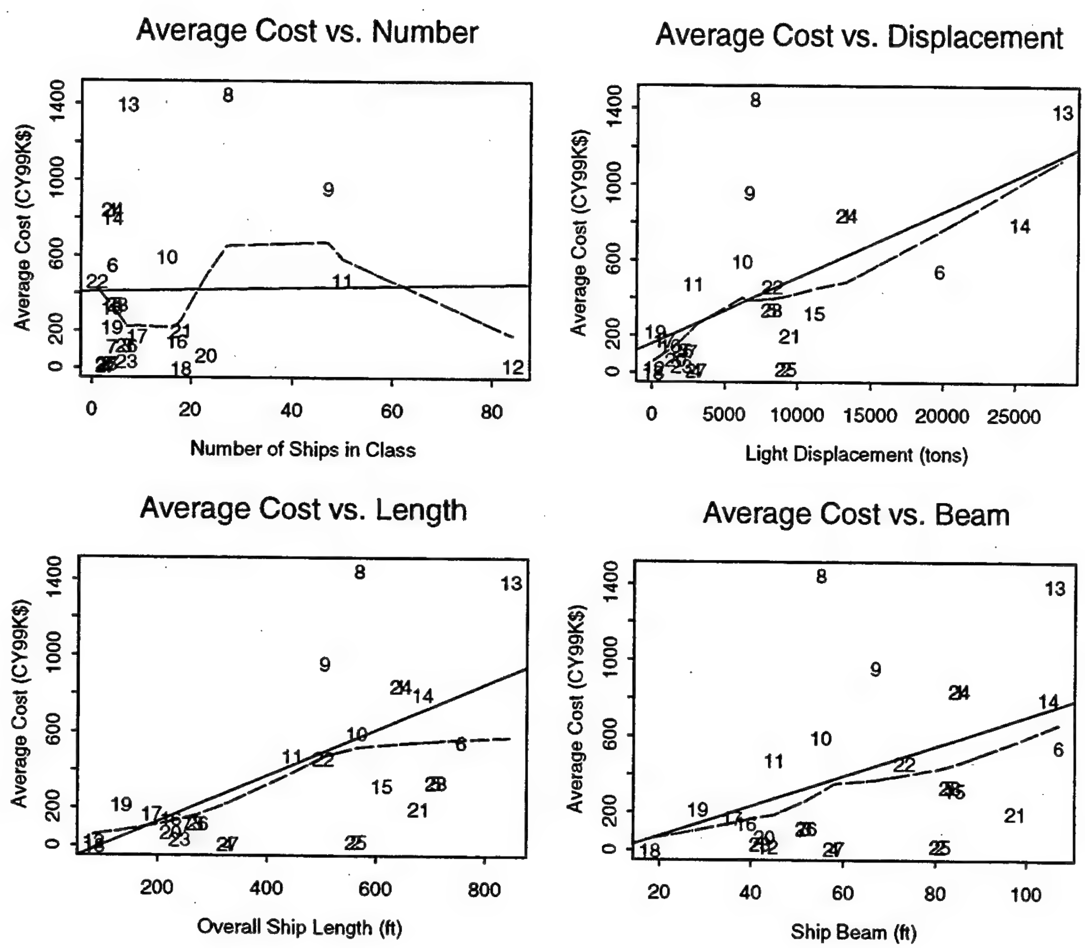


Figure 4: Plot of Average Cost Against Independent Variables.
For number, light displacement, length and beam.

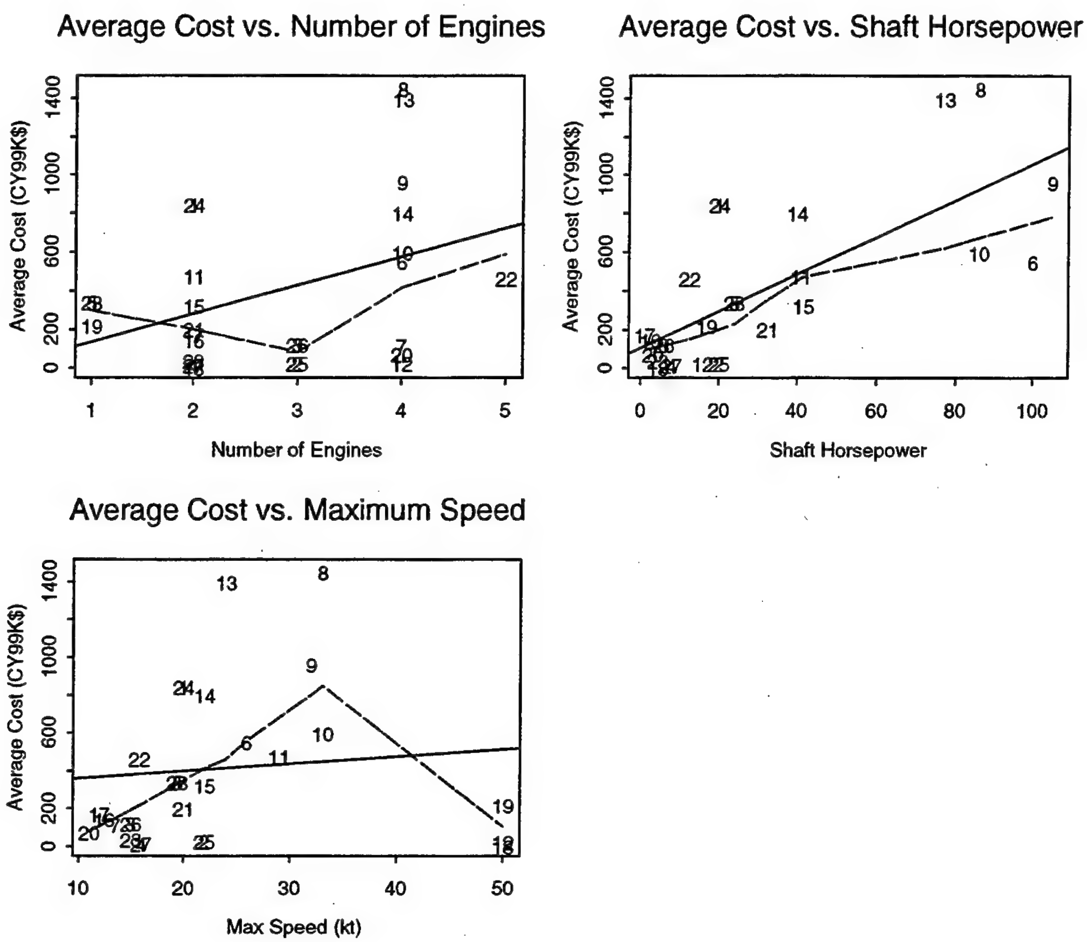


Figure 5: Plot of Average Cost Against Independent Variables.
For number, light displacement, length and beam.

2. Regression Model Generation

This study is intended to generate simple yet accurate models that predict the average cost of a hypothetical system, when only a few system details may be known. Therefore, the model building process shall begin with single variable linear models.

a. Single Variable Linear Models

Because of the manageable number of independent variables, a regression of average cost is performed separately against each independent variable. The models are created using the statistical package S-Plus 4.0. When exploring the performance of each formulation, models are rejected if the indicated significance (or p-value) of the respective F-statistic exceeds 0.2. An example regression summary for the univariate model relating average cost to light displacement is included as Figure 6. A complete summary of models shall be included in Part D, Model Determination and CER Selection.

```
        *** Linear Model ***

Call: lm(formula = AVGCOST ~ DISP, data = shipcost.sm)
Residuals:
    Min     1Q Median    3Q  Max
 -446.7 -198.5 -98.76 145.9 1056

Coefficients:
                  Value Std. Error  t value Pr(>|t|)
   (Intercept) 155.9316  96.0895     1.6228   0.1196
          DISP   0.0353   0.0090     3.9118   0.0008

Residual standard error: 332.8 on 21 degrees of freedom
Multiple R-Squared: 0.4215
F-statistic: 15.3 on 1 and 21 degrees of freedom, the p-value is
0.0008021
```

Figure 6: Sample S-Plus Output for Univariate Model.
Showing a regression of light displacement on *AVGCOST*.

b. Multiple Variable Linear Models

In order to capitalize on the additional information provided by multiple independent variables, S-Plus is also used to create multivariate linear models. The strategy of backward elimination begins with a regression using all possible technical and performance variables. Variables with a significance (t-statistic p-value) >0.2 are removed, one at a time, and a new model is generated. This process continues until all variables appear significant to the model. In Figure 7, the variable *NUM* shall be eliminated because it has a p-value >0.2. The variable *SHP* will not be eliminated during this iteration as only one may be removed at a time. *NUM* shall be removed before *SHP* because in addition to having a p-value >0.2, the sign of the coefficient indicates that as more ships are produced, they become more expensive, an unrealistic characterization.

```
        *** Linear Model ***

Call: lm(formula = AVGCOST ~ NUM + DISP + LEN + BEAM + ENGNUM + SHP, data =
shipcost.sm)
Residuals:
    Min     1Q Median   3Q   Max
 -346.9 -89.88  11.02 94.2 420.2

Coefficients:
                   Value Std. Error   t value  Pr(>|t|)
   (Intercept)  303.4141  221.9271     1.3672    0.1905
           NUM    3.1693    3.0191     1.0497    0.3094
          DISP    0.0442    0.0147     3.0042    0.0084
           LEN    1.9458    0.7230     2.6912    0.0161
          BEAM  -21.3256    6.3467    -3.3601    0.0040
        ENGNUM   54.0394   50.3269     1.0738    0.2989
           SHP    2.2534    2.5713     0.8764    0.3938

Residual standard error: 223.2 on 16 degrees of freedom
Multiple R-Squared: 0.8018
F-statistic: 10.79 on 6 and 16 degrees of freedom, the p-value is 0.00007286
```

Figure 7: Sample S-Plus Output for Multivariate Model.
Showing a regression of six independent variables on *AVGCOST*.

D. MODEL DETERMINATION AND CER SELECTION

Several models performed at a reasonable level of significance. They represent the contenders as CERs. However, the selection of a CER as a useful predictor will also depend on additional information not captured by the p-values.

1. Significant CERs

Significant CERs demonstrate p-values that are less than the limit of 0.2. The insignificant models need not be considered further as useable CERs. Additional statistical measures such as the coefficient of determination and residual standard error shall be used to further evaluate the significant models.

a. Single Variable Models

Five of the seven independent variables demonstrate promise as predictors of acquisition cost. The variables *NUM* and *PROP* have p-values of 0.91 and 0.33 respectively and are removed from further consideration. The remaining formulations, together with the range of variables used in their construction and coefficients of determination and variation, are detailed in Table 7. The range indicates the extreme

Independent Variable	Formulation	Independent Variable Range	R_a^2	*RSE*	*CV*
SHP	*AVGCOST*=103.6+9.545**SHP*	1.16 to 105 khp	0.5490	284.8	68.5%
LEN	*AVGCOST*=-113.2+1.205**LEN*	81 to 844 ft.	0.4122	329.6	79.3%
DISP	*AVGCOST*=155.9+0.0353**DISP*	75 to 28233 tons	0.4014	332.8	80.0%
BEAM	*AVGCOST*=-79.54+7.837**BEAM*	18 to 107 ft.	0.2179	384.3	92.4%
ENGNUM	*AVGCOST*=-12.24+147.0**ENGNUM*	1 to 5 engines	0.1526	401.0	96.4%

Table 7. Single Variable Model Performance.
Showing model formulations, range of independent variables, coefficients of variation and determination and standard error.

values of the independent variable from the data set and is included to specify the scope of values that could credibly be used with the model under the expectation of a linear relationship.

b. Multiple Variable Models

The technique of backward elimination starts with a single model with seven independent variables. By eliminating the insignificant variables, the model was reduced to a four variable model. The regression output is shown in Figure 8. The model displays two shortcomings that must be addressed before accepting the model. First, *BEAM* has a negative coefficient. This would seem to indicate that by making a ship

```
       *** Linear Model ***

Call: lm(formula = AVGCOST ~ DISP + LEN + BEAM + SHP, data =
shipcost.sm)
Residuals:
    Min     1Q Median    3Q   Max
 -354.2 -111.9  -6.38 94.52 477.8

Coefficients:
                   Value Std. Error   t value  Pr(>|t|)
   (Intercept)  477.6136  188.6964    2.5311    0.0209
          DISP    0.0429    0.0142    3.0259    0.0073
           LEN    1.3812    0.6393    2.1606    0.0445
          BEAM  -18.0386    6.0903   -2.9619    0.0083
           SHP    4.7917    2.0396    2.3493    0.0304

Residual standard error: 226.5 on 18 degrees of freedom
Multiple R-Squared: 0.7704
F-statistic: 15.1 on 4 and 18 degrees of freedom, the p-value is
0.00001409

Correlation of Coefficients:
              (Intercept) DISP          LEN       BEAM
          DISP  0.7055
           LEN  0.2221     -0.0070
          BEAM -0.7717     -0.5737       -0.7379
           SHP -0.2347     -0.2184       -0.5646   0.4124
```

Figure 8: Regression Results for Multivariate Model.
Showing a regression of *AVGCOST* on *DISP*, *LEN*, *BEAM* and *SHP*.

fatter, it would be cheaper to build. The second problem is actually the same problem revealed in a new way. The two variables *LEN* and *BEAM* are highly correlated (ρ=-.74), indicating that they both describe similar information. Indeed, in the univariate models, both variables have positive coefficients, indicating that as length or beam is increased, average cost will also increase. The correlation indicates that as a ship gets longer, its beam typically gets larger also. Because of this, some of the price increase due to a larger beam is attributed to the variable coefficient for *LEN*. The negative coefficient for *BEAM* is a correction that reflects the higher cost of narrow ships, compared to wide ones, for a given length.

In an attempt to correct the multicollinearity, the two variables *LEN* and *BEAM* may be combined into the aspect ratio *LENBEAM*, where $LENBEAM = \frac{LEN}{BEAM}$. The backward elimination procedure is then repeated, starting with a full model including all variables except *LEN* and *BEAM*, substituting instead the aspect ratio *LENBEAM*. The resulting three-variable model produces similar results but exhibits much less correlation between variables. The multiple variable models are summarized in Table 8. The regression output is shown in Figure 9.

Independent Variables	Formulation	Independent Variable Range	R_a^2	*RSE*	*CV*
DISP *LEN* *BEAM* *SHP*	*AVGCOST*=477.6+0.0429**DISP*+ 1.381**LEN*+ -18.04**BEAM*+ 4.792**SHP*	*DISP*: 75 to 28233 tons *LEN*: 81 to 844 ft. *BEAM*: 18 to 107 ft. *SHP*: 1.16 to 105 khp	0.5992	226.5	54.5%
DISP *LENBEAM* *ENGNUM*	*AVGCOST* =- 693.8+0.0207**DISP*+ 106.3**LENBEAM*+ 86.63**ENGNUM*	*DISP*: 75 to 28233 tons *LEN*: 81 to 844 ft. *BEAM*: 18 to 107 ft. *ENGNUM*: 1 to 5 engines	0.5607	266.0	64.0%

Table 8. Multiple Variable Model Performance.
Showing model formulations, range of independent variables, coefficients of variation and determination and standard error.

```
*** Linear Model ***

Call: lm(formula = AVGCOST ~ DISP + LENBEAM + ENGNUM, data =
shipcost.sm)
Residuals:
    Min      1Q Median     3Q    Max
 -461.5 -153.4 -81.97 155.5 565.9

Coefficients:
                   Value Std. Error   t value  Pr(>|t|)
   (Intercept) -693.7551  245.0980    -2.8305    0.0107
          DISP    0.0207    0.0083     2.4951    0.0220
       LENBEAM  106.2619   30.9323     3.4353    0.0028
        ENGNUM   86.6332   51.8724     1.6701    0.1113

Residual standard error: 266 on 19 degrees of freedom
Multiple R-Squared: 0.6658
F-statistic: 12.62 on 3 and 19 degrees of freedom, the p-
value is 0.00009065

Correlation of Coefficients:
              (Intercept) DISP          LENBEAM
          DISP  0.2785
       LENBEAM -0.7831     -0.3814
        ENGNUM -0.5880     -0.3377        0.0664
```

Figure 9: Regression Results for Multivariate Model.
Showing a regression of *AVGCOST* on *DISP*, *LENBEAM* and *ENGNUM*.

E. MODEL APPRAISAL AND VALIDATION

Unfortunately, model selection cannot rely simply on statistics. After all, the searching process that determined each model and required significance level of α=0.2

allows useless information to be included into one in every five models. In the end, the models must be individually analyzed for plausibility and their credibility tested.

1. Single Variable Models

The five univariate models may be conveniently clustered into two groups, based upon individual model coefficients of variation and *RSE*. The models for *BEAM* and *ENGNUM* each account for between 16% and 23% of the variability of the database, while *LEN*, *DISP* and *SHP* each describe between 40% and 55% of the variability, as measured by R_a^2.

In addition to outperforming the other two in statistical measures, *LEN*, *DISP* and *SHP* represent information likely to be known or estimable in the uncertain situations for which the models are being developed. Therefore, only the aforementioned three independent variables shall be considered further. Each model shall be referred to by the variable name from which it was formed.

As single variable models, only two regression assumptions are critical to the validity of the model, normality of error terms (residuals) and homoscedasticity. The three models, *LEN*, *DISP* and *SHP*, each demonstrate sufficient adherence to the required assumptions. A graphical summary of each is shown as Figures 10, 11 and 12. A graphical summary of each model is shown in Appendix A.

The *DISP* model has one shortcoming: two points from the database exert significant leverage on the formulation; the points for LHD 1 and LPD 17 play a large role in determining the direction of the OLS line because they are have larger *DISP* values than other points. However, the influence of each point is relatively low,

indicating that the line would have taken a similar form even if the points were not included. As such, all three models are worthwhile CERs for average cost.

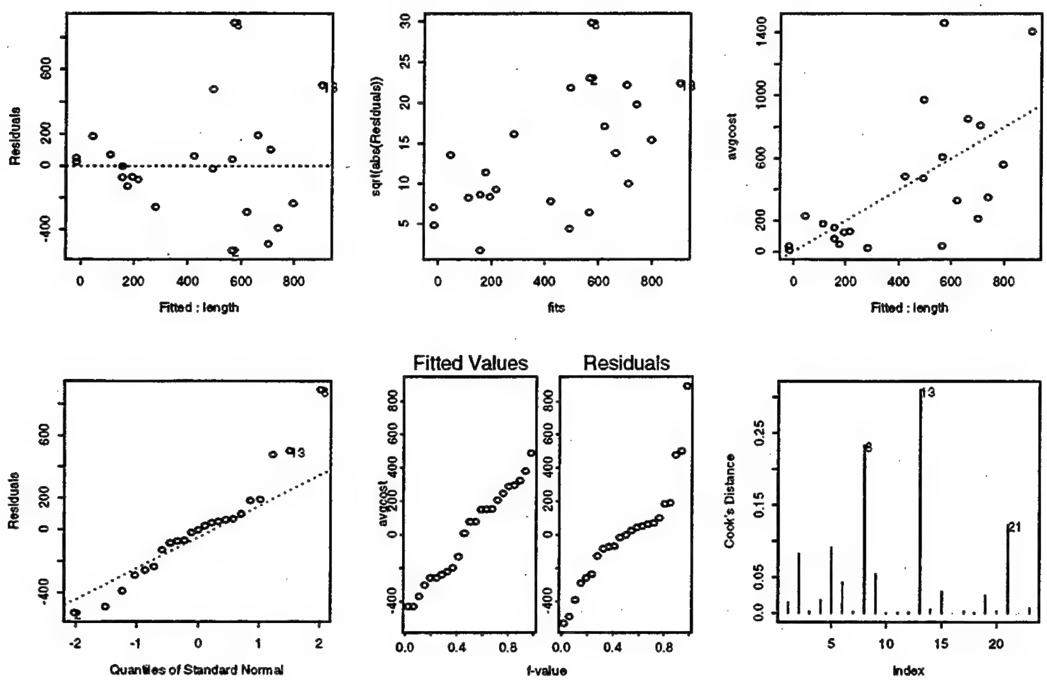


Figure 10: Graphical Performance of the *LEN* Model.
Showing residuals, their absolute values, predicted cost vs. actual cost, a quantile plot and a leverage plot.

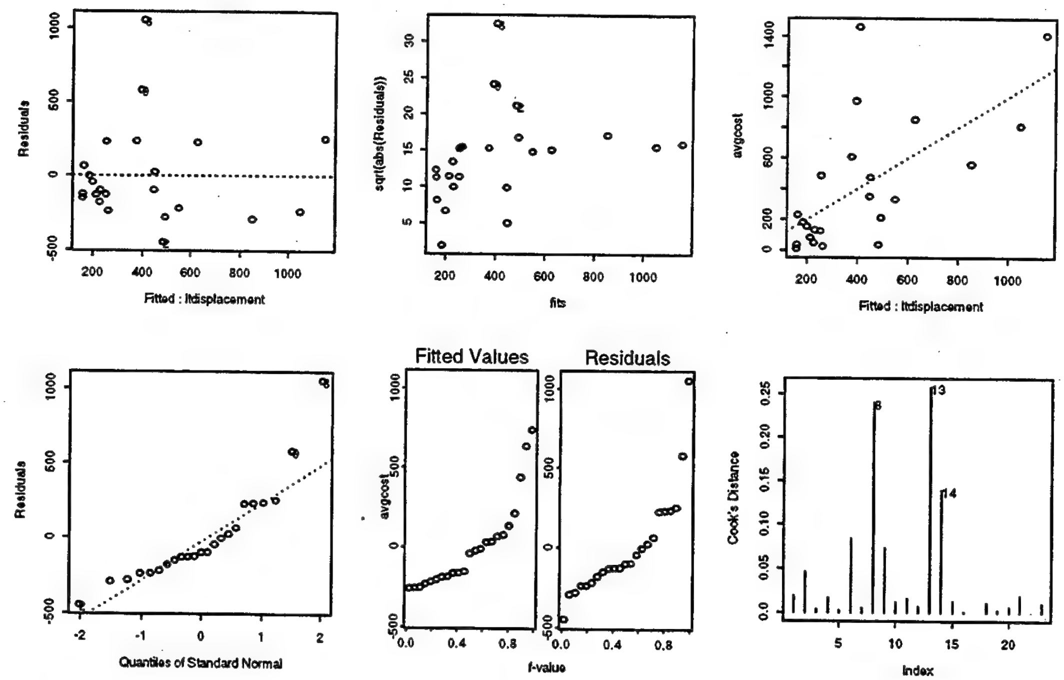


Figure 11: Graphical Performance of the *DISP* Model.
Showing residuals, their absolute values, predicted cost vs. actual cost, a quantile plot and a leverage plot.

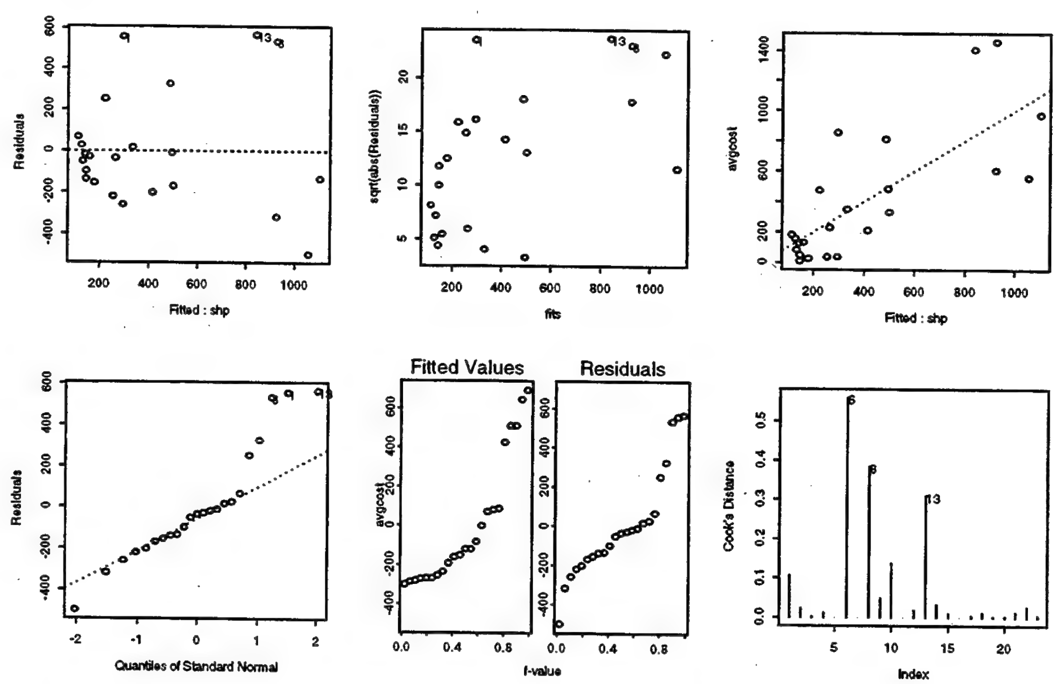


Figure 12: Graphical Performance of the *SHP* Model.
Showing residuals, their absolute values, predicted cost vs. actual cost, a quantile plot and a leverage plot.

2. Multivariate Models

Because referring to the multivariate models by their component independent variables would be difficult, they shall be referred as *MV1*: for the regression of *AVGCOST* on *DISP, LEN, BEAM* and *ENGNUM*; and *MV2*: for the regression of *AVGCOST* on *DISP, LENBEAM* and *SHP*. Both multivariate models share similar statistical and predictive performance. With the exception of the multicollinearity shown with *MV1*, each adequately meets the required regression assumptions. A graphical summary of the multivariate models is included as Figures 13 and 14.

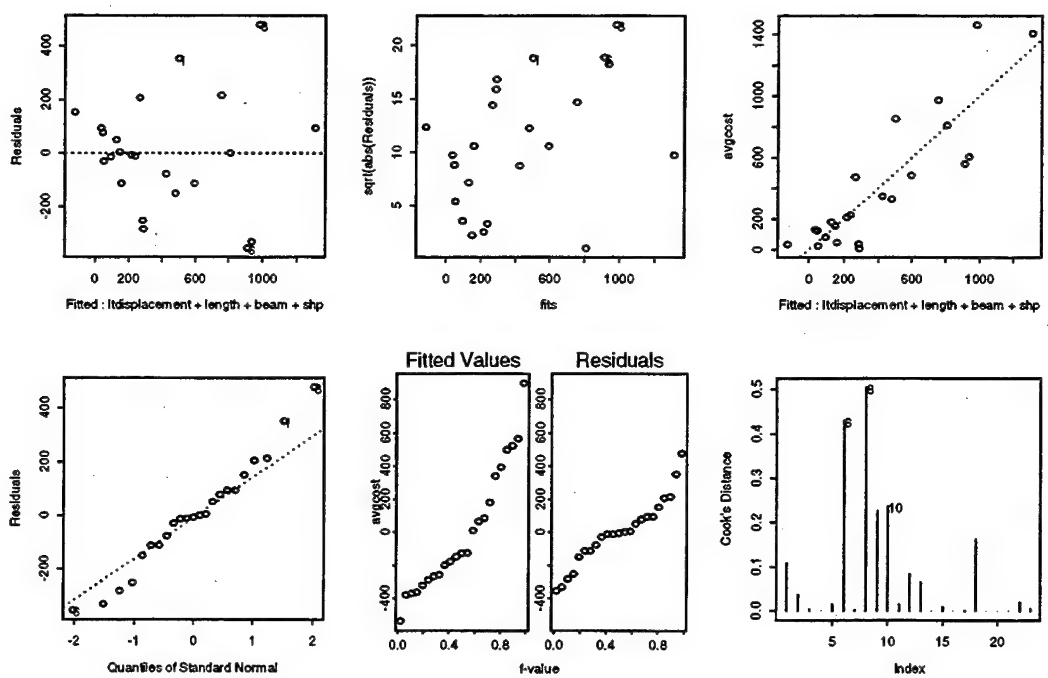


Figure 13: Graphical Performance of the *MV1* Model.
Showing residuals, their absolute values, predicted cost vs. actual cost, a quantile plot and a leverage plot.

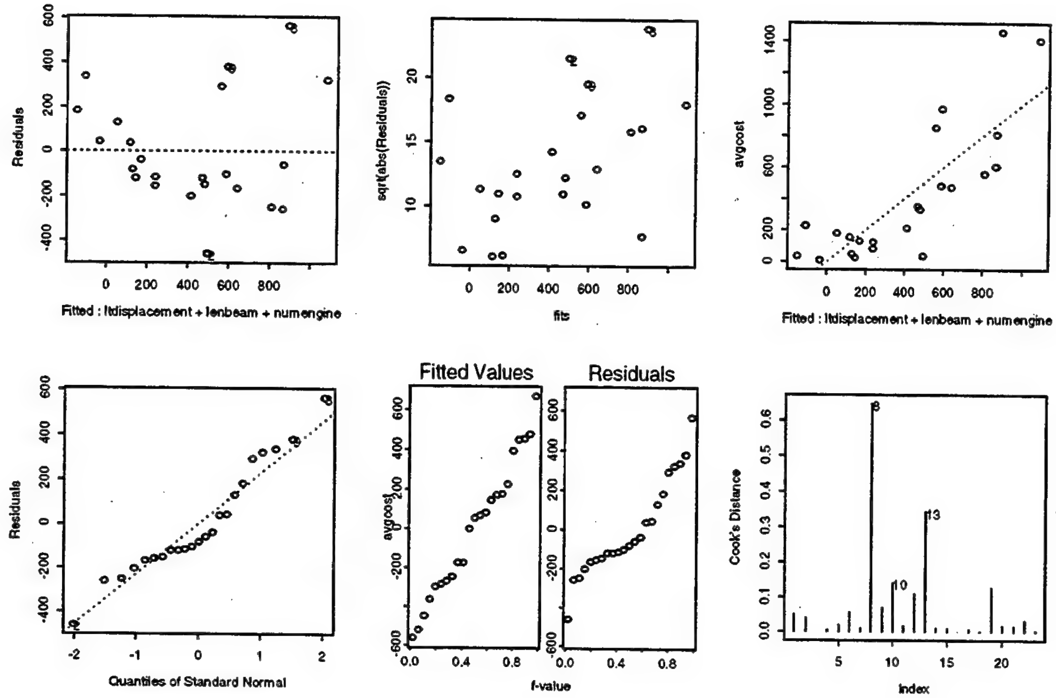


Figure 14: Graphical Performance of the *MV2* Model.
Showing residuals, their absolute values, predicted cost vs. actual cost, a quantile plot and a leverage plot.

Although the four variable *MV1* model exhibits a slightly superior R^2 and *RSE*, the multicollinearity calls into question its output when the highly correlated relationship between *LEN* and *BEAM* fails to hold. Performing a regression of length on beam returns the relationship: $LEN = -79.8 + 8.2 * BEAM$; new vessels that do not approximately follow this relationship will be poorly predicted by *MV1*. In fact, with ρ=-0.74, the *RSE* could be off by 150%. (Hamilton, p. 134-5) Additionally, because both models have nearly equal R_a^2 values, the penalty for having a larger model appears to cancel out the benefits of additional independent variables. As such, *MV2* appears to be the best multivariate model.

3. Model Validation

Since the regression technique has systematically identified the data set deviations and returned a descriptive explanation of them, any model will perform most successfully on the data from which it was generated. Cross-validation must be used to evaluate the single and multivariate models. Because the normalized database includes only twenty-three data points, simple cross-validation shall be used. Although not as powerful as double cross-validation, simple cross-validation still provides insight into how well the model will perform when faced with entirely new data. The statistics generated during cross-validation offer the best characterization of model quality and provide a means of selecting models useful in predicting new values. The $R^2_{a,c}$ values describe the fraction of the variability of a new ship class that should be explained by the models and the RSE_c values provide a realistic standard error for model predictions when used with new data. A summary of the models, their cross-validated coefficients of determination and variation and their standard errors is shown in Table 9. The *SHP* model clearly outperforms the remaining single variable models, rivaling even the multivariate regressions. *MV1* does outperform *MV2*, indicating that the adjustment in calculating R^2_a may penalize *MV1* too harshly. Still, the multicollinearity problems of *MV1* restrict its

use to situations where the collinear relationship between *LEN* and *BEAM* holds, making it less useful as general model.

Model Name	$R^2_{a,c}$	RSE_C	CV_C
DISP	0.2908	344.1	82.7%
LEN	0.2930	343.5	82.6%
SHP	0.4243	310.0	74.5%
MV1	0.4664	276.3	66.4%
MV2	0.3715	308.1	74.1%

Table 9. Cross-validated Model Performance.
Showing model name, cross-validated coefficients of determination and variation and standard error.

IV. RESULTS

This section shall present the models most useful as cost estimating relationships for a parametric cost analysis and discuss the quality of their predictions. The models shall also be used to address a force structure cost analysis as an example application.

A. PRESENTATION OF MODELS

Four models performed reasonably well both in statistical analysis and validation. All predict the average cost of a new ship procurement. However, the models are not sufficient for most cost estimating purposes; their resolution precludes their use in all but rough order of magnitude (ROM) studies. They should be employed only when a ROM answer would meet a study's purpose. The performance of each model shall be summarized and documented below.

1. Summary of Models

The univariate models *LEN, DISP* and *SHP* and the multivariate *MV2* model are each valid for use as parametric cost models. Each performs to a particular level of accuracy. Although either normal or cross-validated statistics could be used to evaluate the models, the cross-validated performance provides a better prediction of how a model will perform when used with entirely new data. As such, the cross-validated statistics should be used when evaluating model suitability for an application and calculating model variability.

a. The LEN Model

The *LEN* model converts the overall length of a ship into an average cost of procurement, in constant 1999 dollars (CY99M$). It is suitable for predicting the average cost of a vessel with an overall length between 81 and 844 feet. A typical prediction may be expected to err by about 83%. The predicted average cost will have a standard error of about $345 million. The model and its performance are summarized in Figure 15.

<u>The *LEN* Model</u>:

$$\boldsymbol{AVGCOST\ (CY99M\$) = -113.23 + 1.2054*LEN}$$

<u>Where</u>:

LEN = Length in ft.

<u>Allowable Range for independent Variable</u>:

Length: 81 to 844 ft.

<u>Cross-validated Statistics</u>:

Adjusted Coefficient of Determination:	29.3%
Coefficient of Variation:	82.6%
Residual Standard Error:	343.5 (CY99M$)

Figure 15: The *LEN* Model Summary.
Predicting average cost with overall length.

b. The DISP Model

The *DISP* model converts the light displacement of a ship into an average cost of procurement, in CY99M$. It is suitable for predicting the average cost of a vessel with a light displacement between 75 and 28233 tons. A typical prediction may be expected to err by about 83%. The predicted average cost will have a standard error of about $345 million. The model and its performance are summarized in Figure 16.

The *DISP* Model:

$$AVGCOST \text{ (CY99M\$)} = 155.93 + 0.0353 * DISP$$

Where:

DISP = Light Displacement in tons.

Allowable Range for independent Variable:

Light Displacement: 75 to 28233 tons.

Cross-validated Statistics:

Adjusted Coefficient of Determination:	29.1%
Coefficient of Variation:	82.7%
Residual Standard Error:	344.1 (CY99M$)

Figure 16: The *DISP* Model Summary.
Predicting overall cost with light displacement.

c. *The SHP Model*

The *SHP* model converts the shaft horsepower (khp) of a ship into an average cost of procurement, in CY99M$. It is suitable for predicting the average cost of a vessel with a propulsion shaft horsepower between 1160 hp and 105,000 hp. A typical prediction may be expected to err by about 75%. The predicted average cost will have a standard error of $310 million. The model and its performance are summarized in Figure 17.

<u>The *SHP* Model</u>:

$$\boldsymbol{AVGCOST\ (\mathrm{CY99M\$}) = 103.63 + 9.5453 * SHP}$$

<u>Where</u>:

SHP = Shaft horsepower in khp.

<u>Allowable Range for independent Variable</u>:

Shaft Horsepower: 1.16 to 105 khp.

<u>Cross-validated Statistics</u>:

Adjusted Coefficient of Determination:	42.4%
Coefficient of Variation:	74.5%
Residual Standard Error:	310.0 (CY99M$)

Figure 17: The *SHP* Model Summary.
Predicting overall cost with propulsion shaft horsepower.

d. The MV2 Model

The *MV2* model converts the overall length, beam, light displacement and number of engines of a ship into an average cost of procurement, in CY99M$. The model may be used for a ship with an overall length between 81 and 844 feet, a beam between 18 and 103 feet, a light displacement between 75 and 28233 tons and with one to five propulsion engines. A typical prediction may be expected to err by about 75%. The predicted average cost will have a standard error of nearly $310 million. The model and its performance are summarized in Figure 18.

The Multivariate *MV2* Model:

$$\boldsymbol{AVGCOST\ (\mathrm{CY99M\$}) = -693.76 + 0.0207{*}DISP + 106.262{*}(LEN/BEAM) + 86.6332{*}ENGNUM}$$

Where:

DISP = Light Displacement in tons.
LEN = Length in ft.
BEAM = Beam in ft.
ENGNUM = Number of Propulsion Engines.

Allowable Range for independent Variable:

Light Displacement: 75 to 28233 tons.
Length: 81 to 844 ft.
Beam: 18 to 103 ft.
Number of Engines: 1 to 5.

Cross-validated Statistics:

Adjusted Coefficient of Determination:	37.2%
Coefficient of Variation:	74.1%
Residual Standard Error:	308.1 (CY99M$)

Figure 18: The Multivariate Model Summary for the *MV2* Model.
Predicting average cost with length, beam, light displacement and shaft horsepower.

All three univariate models, as well as the multivariate model have large coefficients of variation and *RSEs*. Cost estimates may be expected to err by at least 75%, on the average. Again, any user intent on employing these models must be willing to accept predictions that miss the true average procurement cost by a factor of two or more.

2. Model Documentation

A detailed description and documentation of the cost models developed by this study is provided in Appendix B. It is suitable as a stand-alone summary and procedural guide for rough order of magnitude cost models when predicting U.S. Navy conventional surface ship procurement costs. It also contains the necessary uncertainty information to enable cost analysts and decision makers to determine whether the models will be suitable to a particular cost estimating application.

B. ILLUSTRATED EXAMPLE

An example of a suitable use for this type of cost model follows. In this example, a Force Structure Cost Analysis will compare two alternative strategic decisions and the forces necessary to support them. Two competing political strategies requiring different military infrastructures will be investigated. Force compositions are hypothetical; they do not represent strategic concerns of the U.S. Navy, the U.S. Government or any other organization, and are included for illustrative purposes only.

1. The Scenario

Angered by the perception of an increasing percentage of federal funds being devoted to international policies, a prominent domestic special interest group convinces

key political powers to investigate future Naval spending. They argue that the planned programs do not reflect the geopolitical environment but are instead a militaristic holdout from the heady days of nationalism. Congress appoints a panel to look into the matter. The panel identifies two alternatives and proceeds to investigate them. The alternatives represent Naval combatant composition choices:

- **Choice 1:** In support of a strategy that includes nuclear aircraft carriers and carrier battle groups, Choice 1 foresees a four carrier battle group fleet. Each fleet will require the support of the following ships:
  - (2) ACX Advanced Strike Cruisers, a 10000 ton missile cruiser capable of performing extensive strike and anti-air missions.
  - (1) DGX Cooperative Engagement Destroyer, a 500 ft. destroyer designed to leverage expensive sensors from other platforms to perform anti-submarine and shore gunfire support missions.
  - (2) DMX Minesweeping Destroyers, a 550 ft, 6000 ton, 67 ft wide, four engine derivative of current destroyer designs combining anti-submarine and mine-detection missions using remote sensors and active sonar.
  - (1) AFOS Supply and Support ship, a 25000 ton supply ship capable of supporting the remaining vessels.

  In order to support the four battle groups, perform all training and maintenance, and accommodate additional demands, the following fleet configuration is required:
  - (40) ACX
  - (25) DGX
  - (68) DMX
  - (9) AFOS

- **Choice 2:** The alternative strategy forsakes the carrier battle group entirely in favor of a "Jeffersonian" gunboat strategy of identically configured small combatants, dispersed into all regions as global peacekeepers, able to serve as measures of containment while multi-national forces are used in major engagements. The plan focuses on:
  - (200) FGX Multi-mission Frigates, 5000 ton ships equipped with extensive communications capabilities and sufficient defensive weapons to establish a presence in a hostile area, provide extensive reconnaissance information and maintain that presence until multinational forces arrive.

2. The Cost Analysis

While other think tanks evaluate the geopolitical threats of the future and the value of each fleet against the possible threats, you are assigned to provide a cost estimate for each force structure. The information known about each ship is quite modest—detailed studies expected to provide additional information are still years from completion. However, rough cost figures must be provided to the panel to break a deadlock that threatens to stall the passage of the coming budget. The information is sufficient to produce a cost estimate.

a. The Cost of Each Ship and Total Force Structure

The appropriate models will calculate individual average ship cost by entering the known independent variables. The total cost for the fleet types may be generated by multiplying the average cost of each ship by the number required and summing the ship totals. Because each individual ship cost estimate has a normal distribution, the *RSE* of the fleet cost estimates may also be calculated by squaring the sum of the *RSE* values for each ship class, adding the squares, and taking the square root of the sum. As an example, for the Battle Group Support Fleet, the fleet cost estimate $RSE = \sqrt{(40*344.1)^2 + (25*343.5)^2 + (68*308.1)^2 + (9*344.1)^2}$. The models and their results are summarized in Table 10.

Ship Type	Model	Formulation	Average Cost (CY99M$)	Number Required	Total Estimate (CY99M$)	*RSE* (CY99M$)
ACX	DISP	AVGCOST = 155.93 + 0.0353*(10000)	508.93	40	20357.2	344.1
DGX	LEN	AVGCOST = -113.23 + 1.2054*(500)	489.47	25	12236.75	343.5
DMX	MV2	AVGCOST = -693.76 + 0.0207*(6000) + 106.262*(550/67) + 86.6332*(4)	649.27	68	44150.55	308.1
AFOS	DISP	AVGCOST = 155.93 + 0.0353*(25000)	1038.43	9	9345.87	344.1
Total Cost Estimate for Battle Group Support Fleet					86090.37	26517.8
FGX	DISP	AVGCOST = 155.93 + 0.0353*(5000)	332.43	200	66486.0	344.1
Total Cost Estimate for Jeffersonian Gunboat Fleet					66486.0	68820.0

Table 10. Average Cost Estimates for Force Structure Elements.

From the table, the Battle Group Support Fleet should cost approximately $86 billion dollars (CY99) while the Jeffersonian Gunboat Fleet should cost only $66 billion dollars (CY99). Note, however, the *RSE* of each prediction. While the Battle Group Support Fleet estimate could easily vary by a standard deviation, or $26.5 billion dollars, the Jeffersonian Gunboat Fleet estimate could just as easily vary by a standard deviation, or $68.8 billion dollars. A decision maker should be much more confident in the Battle Group Fleet cost estimate than the Jeffersonian Fleet cost estimate.

Because of the OLS assumptions, the error terms, or residuals, are normally distributed. Therefore, additional information may be extracted to help the decision maker analyze alternatives, such as the probabilities of each fleet cost exceeding a particular cost. If a cost of 100 billion dollars will cause the panel to reconsider its decision, you could inform them that despite the cheaper estimated cost of Choice 2, it is more likely to exceed the limit (31.3% vs. 30.0% for Choice 1).

Similarly, if the decision maker wanted to know the probability of Choice 1 exceeding Choice 2, the average (mean) cost of (Choice 1 – Choice 2) and its *RSE* may

be calculated: The average cost is: 86090.37-66486=19604.37(CY99M$). The *RSE* may be calculated likewise:

$$RSE = \sqrt{(40*344.1)^2 + (25*343.5)^2 + (68*308.1)^2 + (9*344.1)^2 + (-200*344.1)^2}$$

Thus, if x = the difference in cost between choices 1 and 2, the probability of Choice 1 exceeding Choice 2 is $P(x \geq 0)$ when x ~ Normal(μ=19604,σ^2=73752). Converting x to a standard normal (Φ), $P(x \geq 0)$ is equivalent to $P\left(\frac{x-19604}{73752} \geq \frac{0-19604}{73752}\right)$ or

$\Phi\left(\frac{-19604}{73752}\right)$, a value of 60.5% from tables of the standard normal, indicating that Choice 1 has a 60.5% probability of costing more than Choice 2.

3. The Conclusion

This example is not designed to champion one fleet structure over another. Instead, it illustrates how a high level model may be used to produce meaningful answers to important questions. Note however the sizable uncertainty associated with the predictions. Although the Jeffersonian fleet is supposed to cost about $67 billion (CY99), with an 87% *CV*, the actual cost could easily be 187%*$67, or $145 billion. If the cost estimating purpose cannot allow such a variation, a different method of cost estimation must be chosen.

V. CONCLUSIONS AND RECOMMENDATIONS

A quick perusal of the models reveals that the estimates they generate are rough indeed. With coefficients of variation between 74% and 83%, actual program costs could easily span from zero to twice the estimate; not an answer on which to stake a reputation. Clearly the models are incapable of making a precise estimate of the average procurement cost of a new Naval ship. However, this interpretation ignores the purpose of the models. An analyst seeking an estimate for the procurement costs of the next six ships for DDG 51 Flight II, a well established and detailed program, would be ill served by using the models produced by this study.

The strength of the four models lies with their minimal data requirements. Because the models are able to turn a single parameter into a cost estimate, they may be used with abstract studies or quick estimates that would defeat a detailed model or estimating process. Still, the coefficients of determination show the models describe less than half of the variability of the data. Assuming that the remaining variation is not merely random error and can actually be predicted, these four models have far to go. One area that appears promising is the inclusion of additional descriptive variables. The physical and performance parameters from the database are able to capture some of the data variability. Other parameters that capture scientific and technical aspects, such as weapons systems and sensor suites, may describe much of the remaining variability. Additional independent variables must be considered.

Additionally, the model results may be leveraged with other general models to provide entire life cycle cost estimates. CERs based only on ship length, displacement or manning are able to estimate yearly operating and support (O&S) costs (Brandt). Results from these O&S models may be leveraged with the results of the procurement cost models to estimate the cost of acquiring and maintaining a particular force structure for its entire effective life. Such an indicator would be a useful measure in determining the life cycle cost or worth of a given force structure.

A. RECOMMENDATIONS

Any parametric method is only as good as the database from which it was created. In order to preserve and hopefully improve the quality of the models, the database must be updated with every new ship or class procured. Fortunately, updating the current CERs requires only the addition of the new data into the database spreadsheet and a new regression.

Additionally, new cost databases offer the promise of models with greater accuracy. A database that detailed acquisition costs by WBS category, especially one that could identify the WBS Level 3 costs under the 'ship' category, i.e. hull structure, propulsion plant, electrical plant, etc. could be used to make models that address only one aspect of the ship's cost. In this way, a model could capitalize on the similarities of several ships without being penalized for the differences. As an example, the AOE 6 and the CG 47 both use similar propulsion systems. Both are also required to keep pace with an aircraft carrier as a mission requirement. A model based solely on propulsion characteristics would probably estimate a similar value for both; a good bet, as both use

the same set of four LM2500 gas turbine engines. However, the model based solely on command and surveillance equipment would likely come to very different estimates for the two classes; the phased array radar and anti-aircraft sensors equipping the CG47 are unlikely to come as cheaply as the sensor suite from the AOE 6.

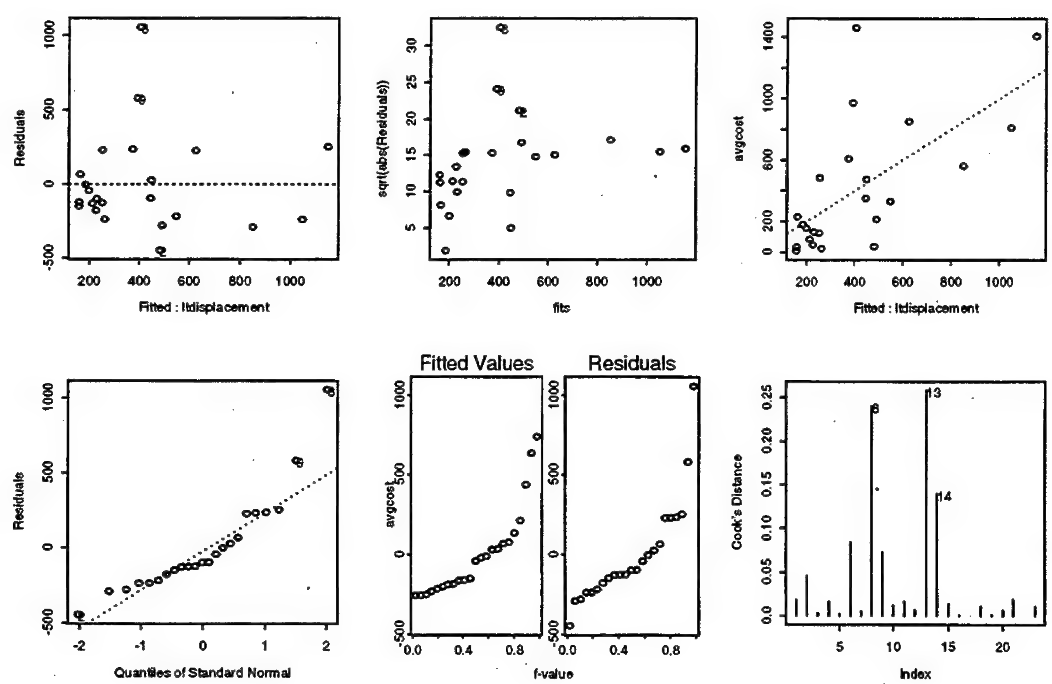
Finally, if cost data can be obtained that details procurement costs in a 'by ship' or 'by lot' format instead of a 'by year' accounting, learning curves could be fitted to the ship cost data. Because the theoretical first unit cost corrects for differences in the number of ships produced, it would allow the cost data to be compared with additional precision. The improvement in accuracy would translate into increased precision of the cost estimate in subsequent analyses and would lower the *RSE* of the models.

Overall, the analyses within this study provide a general-purpose estimator for ship costs when an approximation is sufficient. The sizable *RSE*s of the models prevent them from producing detailed predictions of future program costs, but this point is of little consequence. The models are able to produce a verifiable and defendable estimate from loosely defined parameters when detailed models can not. Within their limited scope, they offer promise as tools able to answer difficult questions in a repeatable, defendable and justifiable manner.

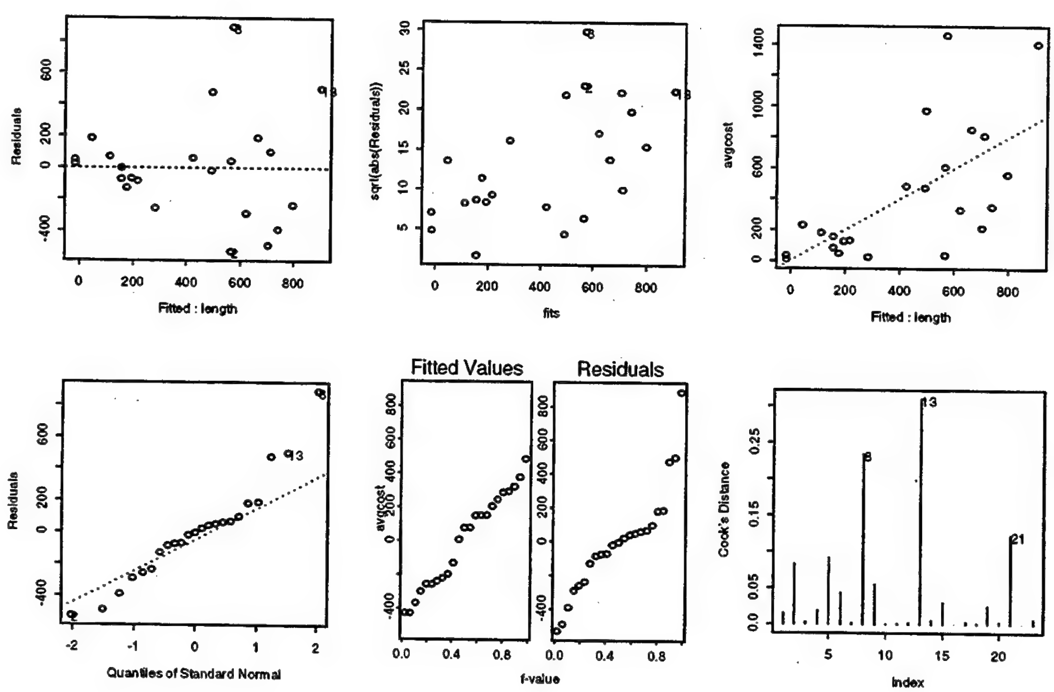
THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. SELECTED MODEL GRAPHICAL PERFORMANCE

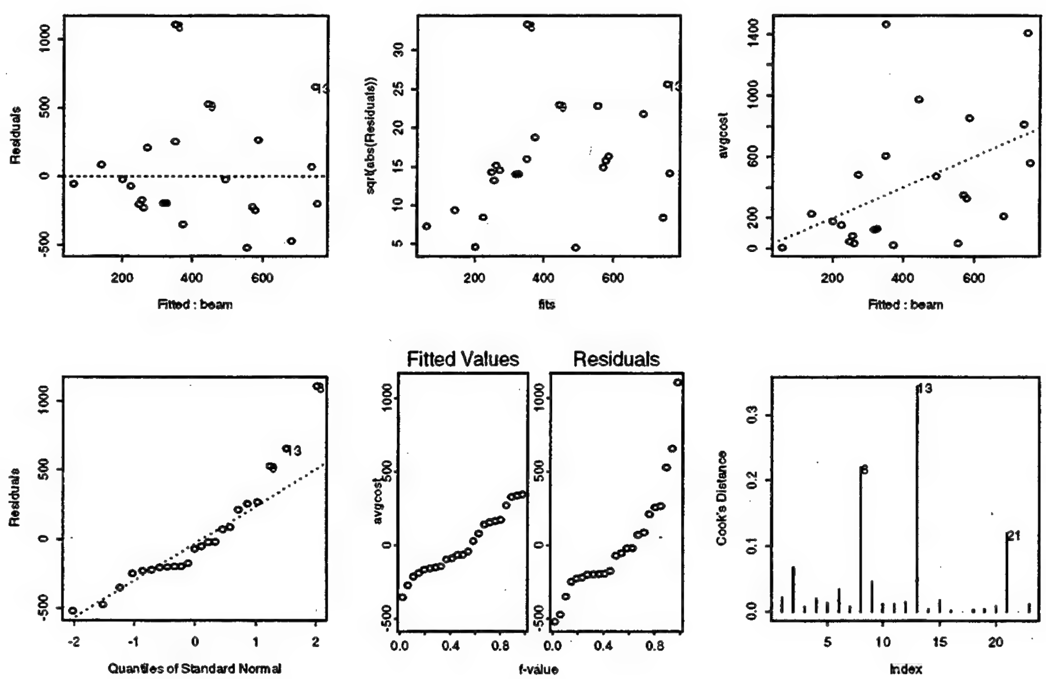
```
Model: Light Displacement

Call: lm(formula = AVGCOST ~ DISP, data = shipcost.sm)
Residuals:
    Min      1Q Median    3Q  Max
 -446.7 -198.5 -98.76 145.9 1056

Coefficients:
                  Value Std. Error  t value Pr(>|t|)
   (Intercept) 155.9316  96.0895     1.6228   0.1196
DISP             0.0353   0.0090     3.9118   0.0008

Residual standard error:       332.8 on 21 degrees of freedom
Multiple R-Squared:            0.4215
Adjusted Multiple R-Squared:   0.4014
F-statistic: 15.3 on 1 and 21 degrees of freedom, the p-value is 0.0008021

Correlation of Coefficients:
              (Intercept)
         DISP     -0.6916

Cross-validated Residual standard error:             344.1
Cross-validated Multiple R-Squared:                  0.3230
Cross-validated Adjusted Multiple R-Squared:         0.2908

Coefficient of Variation:                            80.0%
Cross-Validated Coefficient of Variation:            82.7%
```



```
Model: Length

Call: lm(formula = AVGCOST ~ LEN, data = shipcost.sm)
Residuals:
  Min     1Q Median    3Q Max
 -531 -183.2 -2.858 83.61 889

Coefficients:
                 Value Std. Error  t value  Pr(>|t|)
(Intercept) -113.2234  148.9854   -0.7600    0.4557
        LEN    1.2054    0.3011    4.0028    0.0006

Residual standard error:      329.6 on 21 degrees of freedom
Multiple R-Squared:           0.4328
Adjusted Multiple R-Squared:  0.4122
F-statistic: 16.02 on 1 and 21 degrees of freedom, the p-value is 0.0006454

Correlation of Coefficients:
       (Intercept)
    LEN -0.8872

Cross-validated Residual standard error:                343.5
Cross-validated Multiple R-Squared:                     0.3251
Cross-validated Adjusted Multiple R-Squared:            0.2930

Coefficient of Variation:                               79.3%
Cross-Validated Coefficient of Variation:               82.6%
```

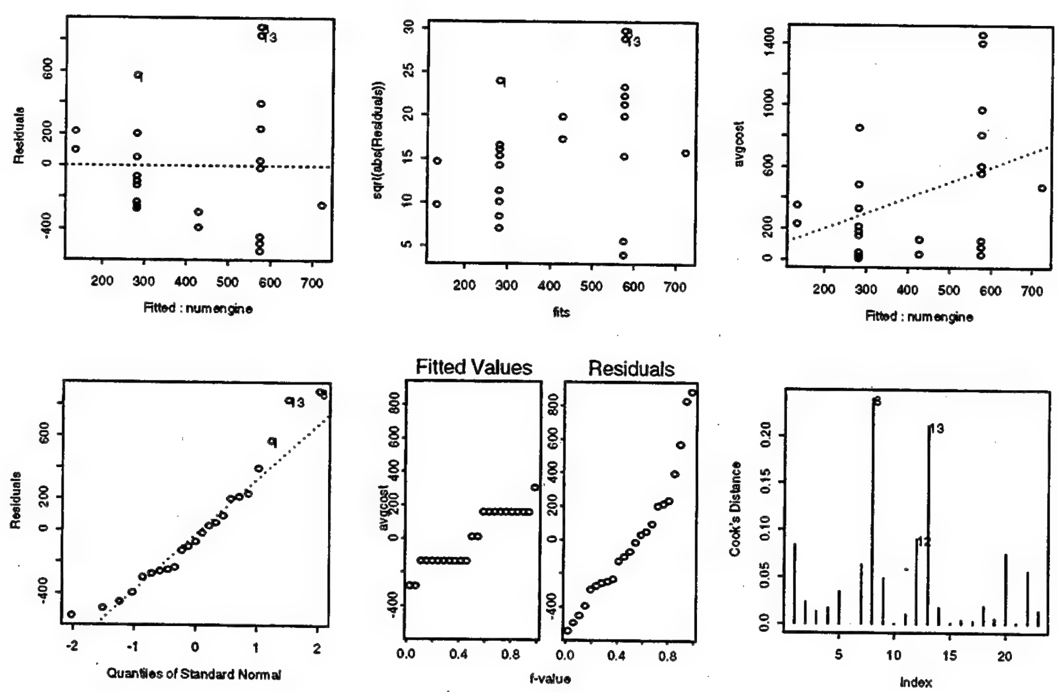


```
Model: Beam

Call: lm(formula = AVGCOST ~ BEAM, data = shipcost.sm)
Residuals:
    Min    1Q Median  3Q  Max
 -519.9 -212 -71.01 149 1108

Coefficients:
                Value Std. Error  t value  Pr(>|t|)
(Intercept)  -79.5442  214.0376   -0.3716    0.7139
       BEAM    7.8365    3.1393    2.4963    0.0209

Residual standard error:      384.3 on 21 degrees of freedom
Multiple R-Squared:           0.2288
Adjusted Multiple R-Squared:  0.2179
F-statistic: 6.231 on 1 and 21 degrees of freedom, the p-value is 0.02095

Correlation of Coefficients:
     (Intercept)
BEAM -0.9273

Cross-validated Residual standard error:              398.6
Cross-validated Multiple R-Squared:                   0.0914
Cross-validated Adjusted Multiple R-Squared:          0.0481

Coefficient of Variation:                             92.4%
Cross-Validated Coefficient of Variation:             95.8%
```

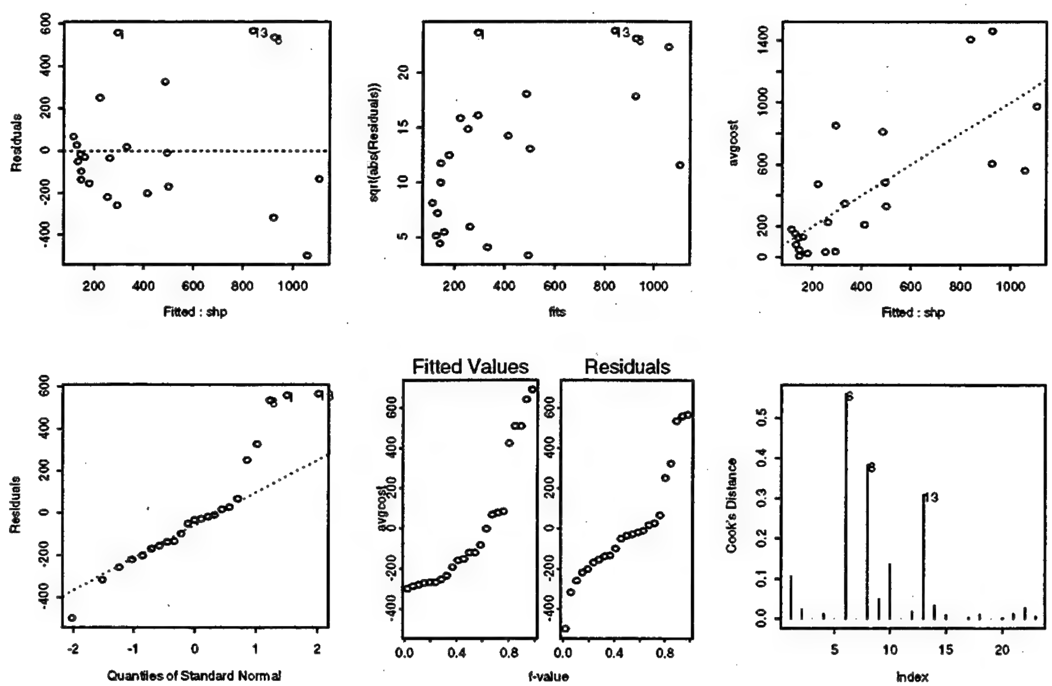


Model: Number of Engines

```
Call: lm(formula = AVGCOST ~ ENGNUM, data = shipcost.sm)
Residuals:
    Min     1Q Median    3Q   Max
 -541.5 -265.2  -70.3 208.4 883.6

Coefficients:
                 Value Std. Error  t value  Pr(>|t|)
(Intercept)   -12.2351  229.6810   -0.0533    0.9580
     ENGNUM   146.9683   73.4341    2.0014    0.0584

Residual standard error:       401 on 21 degrees of freedom
Multiple R-Squared:            0.1602
Adjusted Multiple R-Squared:   0.1526
F-statistic: 4.005 on 1 and 21 degrees of freedom, the p-value is 0.05844

Correlation of Coefficients:
          (Intercept)
ENGNUM        -0.9314

Cross-validated Residual standard error:              417.6
Cross-validated Multiple R-Squared:                   0.0028
Cross-validated Adjusted Multiple R-Squared:         -0.0447

Coefficient of Variation:                             96.4%
Cross-Validated Coefficient of Variation:            100.4%
```



```
Model: Shaft Horsepower

Call: lm(formula = AVGCOST ~ SHP, data = shipcost.sm)
Residuals:
    Min      1Q Median    3Q   Max
 -498.7 -162.8 -35.24 45.93 566.4

Coefficients:
               Value Std. Error  t value Pr(>|t|)
(Intercept) 103.6345  83.3011     1.2441   0.2272
        SHP   9.5453   1.7856     5.3457   0.0000

Residual standard error:      284.8 on 21 degrees of freedom
Multiple R-Squared:           0.5764
Adjusted Multiple R-Squared:  0.5490
F-statistic: 28.58 on 1 and 21 degrees of freedom, the p-value is 0.00002662

Correlation of Coefficients:
    (Intercept)
SHP -0.7012

Cross-validated Residual standard error:              310.0
Cross-validated Multiple R-Squared:                   0.4504
Cross-validated Adjusted Multiple R-Squared:          0.4243

Coefficient of Variation:                             68.5%
Cross-Validated Coefficient of Variation:             74.5%
```



Model: Multivariate ONE (Light Displacement, Length, Beam, Shaft Horsepower)

Call: lm(formula = AVGCOST ~ DISP + LEN + BEAM + SHP, data = shipcost.sm)

Residuals:

Min	1Q	Median	3Q	Max
-354.2	-111.9	-6.38	94.52	477.8

Coefficients:

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	477.6136	188.6964	2.5311	0.0209
DISP	0.0429	0.0142	3.0259	0.0073
LEN	1.3812	0.6393	2.1606	0.0445
BEAM	-18.0386	6.0903	-2.9619	0.0083
SHP	4.7917	2.0396	2.3493	0.0304

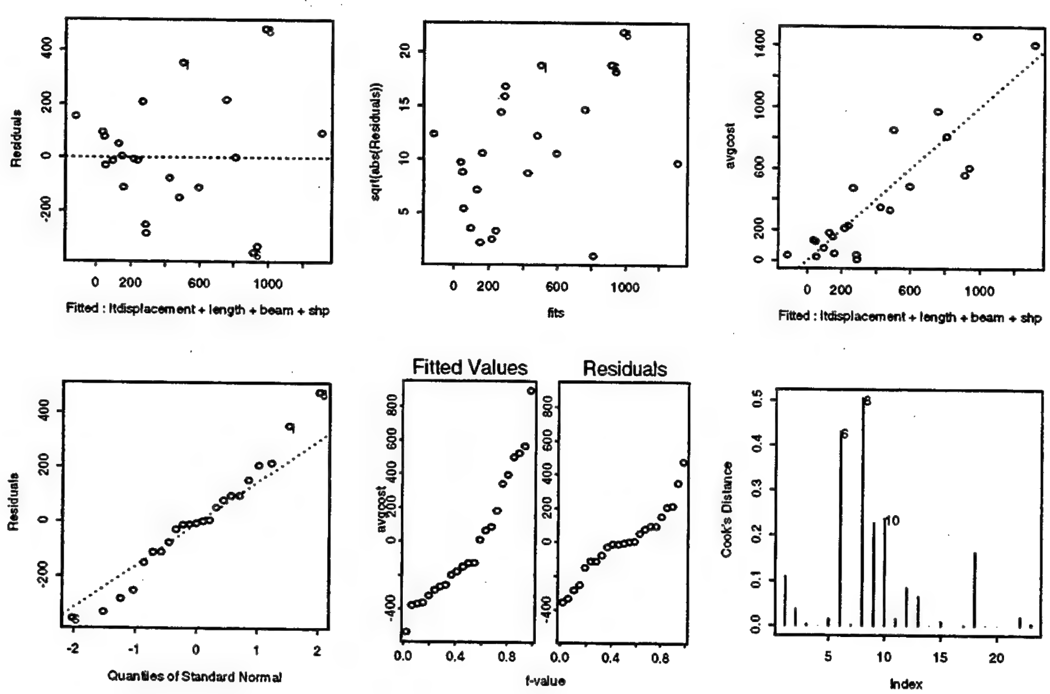
Residual standard error: 226.5 on 18 degrees of freedom
Multiple R-Squared: 0.7704
Adjusted Multiple R-Squared: 0.5992
F-statistic: 15.1 on 4 and 18 degrees of freedom, the p-value is 0.00001409

Correlation of Coefficients:

	(Intercept)	DISP	LEN	BEAM
DISP	0.7055			
LEN	0.2221	-0.0070		
BEAM	-0.7717	-0.5737	-0.7379	
SHP	-0.2347	-0.2184	-0.5646	0.4124

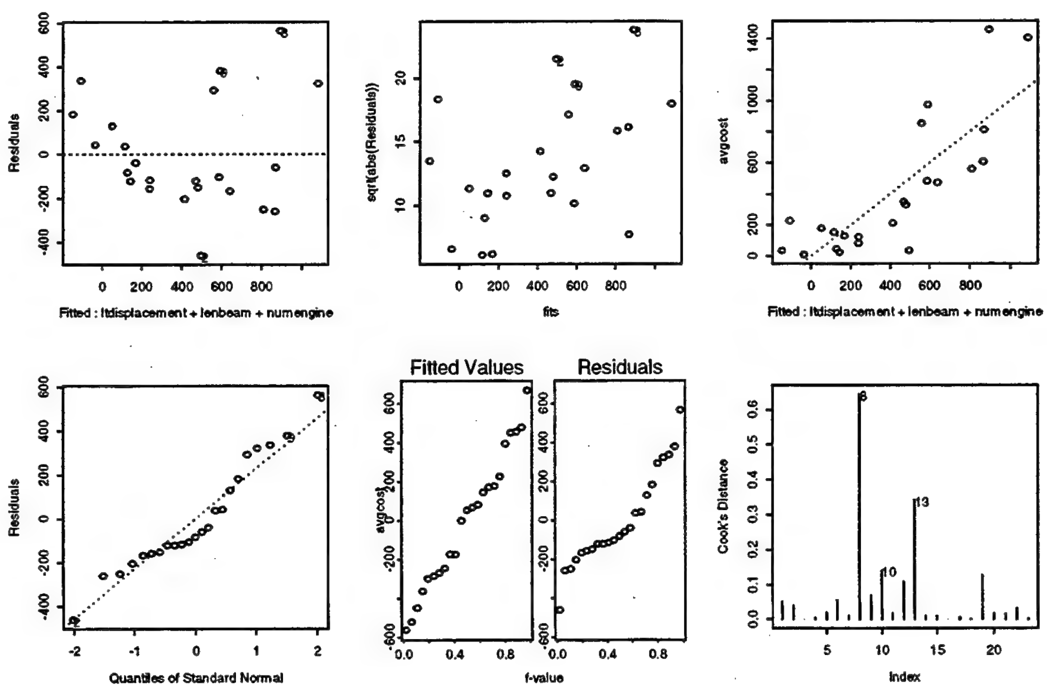
Cross-validated Residual standard error: 276.3
Cross-validated Multiple R-Squared: 0.5634
Cross-validated Adjusted Multiple R-Squared: 0.4664

Coefficient of Variation: 54.5%
Cross-Validated Coefficient of Variation: 66.4%



Model: Multivariate TWO (Light Displacement, Length/Beam, Number of Engines)

```
Call: lm(formula = AVGCOST ~ DISP + LENBEAM + ENGNUM, data = shipcost.sm)
Residuals:
    Min     1Q Median    3Q   Max
 -461.5 -153.4 -81.97 155.5 565.9

Coefficients:
                  Value Std. Error  t value  Pr(>|t|)
   (Intercept) -693.7551  245.0980  -2.8305   0.0107
          DISP    0.0207    0.0083   2.4951   0.0220
       LENBEAM  106.2619   30.9323   3.4353   0.0028
        ENGNUM   86.6332   51.8724   1.6701   0.1113

Residual standard error:       266 on 19 degrees of freedom
Multiple R-Squared:            0.6658
Adjusted Multiple R-Squared:   0.5607
F-statistic: 12.62 on 3 and 19 degrees of freedom, the p-value is 0.00009065

Correlation of Coefficients:
             (Intercept) DISP          LENBEAM
        DISP  0.2785
     LENBEAM -0.7831      -0.3814
      ENGNUM -0.5880      -0.3377        0.0664

Cross-validated Residual standard error:             308.1
Cross-validated Multiple R-Squared:                  0.4572
Cross-validated Adjusted Multiple R-Squared:         0.3715

Coefficient of Variation:                            64.0%
Cross-Validated Coefficient of Variation:            74.1%
```



THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. DOCUMENTATION OF THE PARAMETRIC COST MODEL

Title: **Top-Level U.S. Navy Conventional Surface Ship Parametric Procurement Cost Model**

Purpose: To estimate average procurement costs for conventional U.S. Navy surface ships using one of the following three physical parameters: ship overall length, ship light displacement or propulsion shaft horsepower; or the four physical parameters: ship overall length, ship beam, ship light displacement and number of engines.

Applicability: This top-level procurement cost model is a parametric cost-estimating tool which will provide cost analysts and decision makers with a standardized method for calculating ship procurement cost estimates, based upon historical data, for U.S. Navy conventional ships (excluding nuclear aircraft carriers and submarines). It may be used to estimate costs of roughly defined ships when a significant uncertainty in the estimate is acceptable, such as Rough Order of Magnitude estimates and Force Structure Cost Analyses.

Model Description: This top-level procurement cost model consists of three univariate cost estimating relationship (CER) equations and one multivariate CER. All CERs predict average ship procurement costs in constant year 1999 dollars. The first univariate CER uses ship overall length in feet, the second univariate CER uses ship light displacement in tons and the third univariate CER uses ship propulsion shaft horsepower in thousands of horsepower. (khp) The multivariate CER uses light displacement in tons, the ratio of length in feet to beam in feet and number of propulsion engines. All four CERs were developed using a historical cost database representing major ship acquisition programs from 1973 to present, including frigates, destroyers, cruisers, amphibious assault ships, landing ship docks, oilers, fast combat support ships, combat stores ships, hydrofoils, air-cushion vehicles, oceanographic research ships, tugs, cable repair ships and minesweepers.

Status/Availability:	The top-level procurement cost models are complete. Periodic updates of historical data are strongly recommended. The original release date for this cost model is tentatively scheduled for the first quarter of CY2000. The models may be adapted for use in spreadsheets for ease of calculation and presentation.
Input Variables: (including range)	Ship overall length (ft.) (81-844) or Ship light displacement (tons) (75-28233) or Ship propulsion shaft horsepower (khp) (1.16-105) or Ship overall length (ft.) (81-844) and Ship beam (ft.) (18-103) and Ship light displacement (tons) (75-28233) and Ship number of propulsion engines (1-5)
Output:	Average cost values in constant 1999 (CY99M$) dollars bounded by the residual standard error of the CER model in CY99M$.
Data Sources:	Cost data was compiled from *U.S. Weapon Systems Costs*, Data Search Associates (1999,1995,1990,1987), by Ted Nicholas and Rita Rossi Performance and technical data was compiled from JANE's Fighting Ships, JANEs Publishing, Inc. (1998-99, 1995-96, 1990-91, 1984-85)
Point of Contact:	LCDR Timothy P. Anderson Department of Operations Research Naval Postgraduate School, Monterey, CA
Ground Rules/ Assumptions/ Limitations:	Nuclear powered vessels and submarines were removed from the database in order to normalize data. All data was normalized to CY99M$.
Software:	The CER equations may be employed with any spreadsheet or programming language.

CER Equations:	AVGCOST = -113.23 + 1.2054*LEN	RSE=343.5
	AVGCOST = 155.93 + 0.0353*DISP	RSE=344.1
	AVGCOST = 103.63 + 9.5453*SHP	RSE=310.0
	AVGCOST = -693.76 + 0.0207*DISP + 106.262*(LEN/BEAM) + 86.6332*ENGNUM	RSE=308.1

Validation: Validation was conducted using the historical database and the technique of simple cross-validation. Standard errors reported for the models are the cross-validated estimates instead of the RSEs generated by the regression. The larger magnitude of the cross-validated RSEs reflects the additional uncertainty of predicting new data with the models.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Brandt, James M. *A Parametric Cost Model for Estimating Operating and Support Costs of U.S. Navy (Non-nuclear) Surface Ships.* Master's Thesis, Operations Research Department, Naval Postgraduate School, Monterey, CA. 1999.

Data Search Associates, *U.S. Weapon System Costs*, Fountain Valley, CA: Data Search Associates. 1999, 1995, 1990, 1987.

Department of Defense, *Joint Warfighting Science and Technology Plan* (1997). Department of Defense.

Department of Defense Handbook MIL-HDBK-881, *Work Breakdown Struc*ture (1998). Department of Defense. Available: http://www.acq.osd.mil/pm/newpolicy/wbs/mil_hdbk_881/mil_hdbk_881.htm.

Department of Defense Instruction 5000.2R Change 3, *Mandatory Procedures for MDAPs and MAIS Acquisition Programs* (1998), Department of Defense.

Devore, J.L., (1995). *Probability and Statistics for Engineering and the Sciences* (4th). Pacific Grove, CA: Duxbury Press.

Hamilton, L.C. (1992). *Regression with Graphics, a Second Course in Applied Statistics.* Belmont, CA: Wadsworth Publishing Co., Inc.

Jane's Publishing, Inc. *JANE'S fighting Ships.* New York: Author, 1998-99, 1995-96, 1990-91, 1984-85.

Mosteller, F., Tukey, J.W. (1977). *Data Analysis and Regression.* Reading, MA: Addison-Wesley Publishing.

Naval Air Warfare Center (NAWCWPNS*) Weapon Systems Cost Analysis Course Handbook.* China Lake, CA: Author, 1992.

[OA4702]. Unpublished curriculum text for Naval Postgraduate School Course OA-4702: Cost Estimation. Operations Research Department, Naval Postgraduate School, 1999.

Peters, Katherine M. (1999). *Future Fleet.* Government Executive Magazine [On-line]. Available: http://www.govexec.com/features/0999/0999s3.htm.

Polmar, Norman. *Naval Institute Guide to Ships and Aircraft of the U.S. Fleet* (16th). Annapolis, MD: Naval Institute Press (1997).

Publishing Manual of the American Psychological Association (3rd). (1984). Washington, D.C.: American Psychological Association.

Scott, R, Spiker, R., & Thibault, M. (1995). *Parametric Cost Estimating Handbook.* (1995). Washington, D.C.: Joint Government/Industry Parametric Cost Estimating Initiative Steering Committee, Department of Defense.

SPLUS Guide to Statistics (v4.0) (1997) Seattle, WA: Mathsoft.

SPLUS Programmer's Guide (v4.0) (1997) Seattle, WA: Mathsoft.

SPLUS User's Guide (v4.0) (1997) Seattle, WA: Mathsoft.

Transportation Institute. *Maritime Glossary* [On-line] 1999. Available: http://www.trans-inst.org/.

United States Army Logistics Management College (USALMC). *Cost Estimating Reference Book.* Fort Lee, VA: Author, 1991.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
   8725 John J. Kingman Rd., STE 0944
   Fort Belvoir, Virginia 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, California 93943-5101

3. LCDR Timothy P. Anderson, Code OR/At 2
   Department of Operations Research
   Naval Postgraduate School
   Monterey, California 93943-5000

4. Professor James N. Eagle, Code OR/Er 1
   Department of Operations Research
   Naval Postgraduate School
   Monterey, California 93943-5000

5. Naval Sea Systems Command 3
   NAVSEA 017, NC3/11S08
   Attn: Lisa Pfeiffer
   2531 Jefferson Davis Hwy
   Arlington VA 22242

6. CDR William Valentine N816B 3
   2000 Navy Pentagon
   Washington, DC 20350-2000

7. Director, Naval Center for Cost Analysis 3
   3801 Nebraska Ave, NW
   Washington, DC 20393-5444

8. LT Kirk J. Loftus 3
   2741 Mt. Pleasant Rd.
   Kelso, Washington 98626